

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 25, 2002, 20:25:44 ; Search time 312.78 Seconds
(without alignments)
4138.863 Million cell updates/sec

Title: US-09-509-779-3

Perfect score: 754
Sequence: 1 ATGGCCGACGTGAGACGCG.....CAATTAAGTCAGTTTAA 754

Scoring table: IDENTITY_NNC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N.Geneseq_032802:*

- 1: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
- 2: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
- 3: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
- 4: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
- 5: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
- 6: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
- 7: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
- 8: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
- 9: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
- 10: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
- 11: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
- 12: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
- 13: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
- 14: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
- 15: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
- 16: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
- 17: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
- 18: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
- 19: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
- 20: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
- 21: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
- 22: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
- 23: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
- 24: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	754	100.0	754	20	AAx87314
2	752.4	99.8	754	20	AAx87317
3	752.4	99.8	754	20	AAx87318
4	752.4	99.8	754	20	AAx87320
5	752.4	99.8	754	20	AAx87321
6	752.4	99.8	754	20	AAx87324
7	752.4	99.8	754	20	AAx87325
8	752.4	99.8	754	20	AAx87327
9	752.4	99.8	754	20	AAx87328

10	752.4	99.8	754	20	AAx87331	Human sensitive to
11	750.8	99.6	754	20	AAx87319	Human sensitive to
12	750.8	99.6	754	20	AAx87322	Human sensitive to
13	750.8	99.6	754	20	AAx87323	Human sensitive to
14	750.8	99.6	754	20	AAx87329	Human sensitive to
15	750.8	99.6	754	20	AAx87330	Human sensitive to
16	747.6	99.2	754	20	AAx87326	Human sensitive to
17	730	96.8	747	20	AAx87315	Human sensitive to
18	689.4	91.4	836	22	AAx87315	Human sensitive to
19	688.4	91.4	836	22	AAx87319	Human sensitive to
20	689.4	91.4	836	22	AAx87304	Human sensitive to
21	648	85.9	706	20	AAx87316	Human sensitive to
22	501.8	66.6	1152	22	AAx87347	Human sensitive to
23	344.2	45.6	441	21	AAx87493	Human sensitive to
24	337.2	44.7	342	21	AAx86883	Human sensitive to
25	337.2	44.7	342	22	AAx83570	Human sensitive to
26	334	44.3	1140	20	AAx87313	Human sensitive to
27	326	43.2	596	22	AAx87502	Human sensitive to
28	315.8	41.9	389	21	AAx87944	Human sensitive to
29	315.8	41.9	389	22	AAx86882	Human sensitive to
30	314.4	41.7	388	21	AAx87978	Human sensitive to
31	314.4	41.7	388	22	AAx87716	Human sensitive to
32	197	26.1	224	20	AAx84085	Human sensitive to
33	181.2	24.0	630	22	AAx82239	Human sensitive to
34	168	22.3	264	22	AAx86209	Human sensitive to
35	114	15.1	169	20	AAx84482	Human sensitive to
36	86.4	11.5	416	22	AAx83369	Human sensitive to
37	83	11.0	498	23	AAx84673	Human sensitive to
38	83	11.0	10025	23	AAx84672	Human sensitive to
39	81	10.7	557	22	AAx87861	Human sensitive to
40	80.2	10.6	692	23	AAx82527	Human sensitive to
41	77.8	10.3	5347	24	AAx84844	Human sensitive to
42	76	10.1	327	21	AAx86882	Human sensitive to
43	76	10.1	482	21	AAx83896	Human sensitive to
44	76	10.1	508	21	AAx84978	Human sensitive to
45	75.2	10.0	504	21	AAx84980	Human sensitive to

ALIGNMENTS

RESULT 1	
AAx87314	standard; cDNA; 754 BP.
ID	AAx87314
AC	AAx87314;
DT	27-SEP-1999 (first entry)
DE	Human sensitive to apoptosis (SAG) gene.
XX	
XX	
KW	SAG gene; sensitive to apoptosis; human; cancer; tumour;
KW	neurodegenerative disease; muscular dystrophy; wound healing;
KW	vulnerable; therapy; ds.
OS	Homo sapiens.
XX	
XX	
FT	Key
FT	CDS
FT	Location/Qualifiers
FT	1.342
FT	/tag- a
PD	W09932514-A2.
PD	01-JUL-1999.
XX	
XX	
PF	15-DEC-1998; 98WO-US26705.
XX	
PR	11-SEP-1998; 98US-0099840.
PR	19-DEC-1997; 97US-0068179.
XX	
PA	(WARN) WARNER LAMBERT CO.
XX	
PI	Sun Y;

XX WPI: 1999-430152/36.
DR P-PSDB: AAY06492.

PT SAG: Sensitive to Apoptosis Gene and related proteins, useful for
promoting cell growth and protecting cells against apoptosis

PS Claim 9; Page 50-51; 84pp; English.

XX This is the nucleotide sequence of human sensitive to apoptosis
CC (SAG) cDNA, which codes for a novel redox-sensitive, haem-binding
CC protein (see AAY06492) with a zinc RING finger domain that promotes
CC cell growth, protects cells from apoptosis, scavenges oxygen
CC radicals and can be used for the reversion of a tumour phenotype.
CC The cDNA was isolated from a HeLa cell cDNA library using mouse
CC SAG cDNA (see AAX87313) as probe. SAG is highly conserved among
CC species. Disruption in yeast was shown to be lethal. SAG deletion
CC mutants (see AAX87315-16) have been identified in human cancer lines,
CC suggesting a role in carcinogenesis. SAG genes, and mutant SAG
CC genes, can be used to protect cells from apoptosis induced by redox
CC reagents. Antisense SAG genes can be used to inhibit the growth of
CC tumour cells. The SAG genes can also be used for the recombinant
CC production of the SAG proteins. The SAG proteins can be used to
CC scavenge oxygen radicals in organisms and to promote wound healing.
CC Additionally, the SAG genes or their complements can be used to
CC promote or inhibit the growth of plant cells (all claimed). The SAG
CC protein is also an ideal molecular target in the development of
CC drugs against neurodegenerative disorders, cancers and muscle
CC dystrophy.

SO Sequence 754 BP; 205 A; 155 C; 201 G; 193 T; 0 other;

Query Match 100.0%; Score 754; DB 20; Length 754;
Best Local Similarity 100.0%; Pred. No. 2.2e-231;
Matches 754; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCCGACCTGGAGACGAGAGAAACCTGGCCCTGACCTCAGCCGAGATGCA 60
DB 1 atgcccgaactggaagacggaagaaacctgcccctcctcactccggagactca 60
QY 61 GGCCTCAAGTCGGAGGCGCAAGATGTTCTCCCTCAAGAAAGTGAACGGCGGCATG 120
DB 61 ggcctcaagtcggagggcgcaagatggttccctcaagaagtggaacggcggcacatg 120
QY 121 TGAAGCTGGAGACGTGAGTCCGATACGTGCGCCATCTGCAAGGTCCAGTGATGATGCC 180
DB 121 tgaagctggagacgtgagtcgatacgtgcacatctgcaggtccagtgatgatgcc 180
QY 181 TGTCTTAGTGAAGTCAAGCTGAAACAAACAGAGAGACGTGTGTCGTGGGAGAAATGT 240
DB 181 tgtcttagtgaagtcaagctgaacaaacaaagagactgtgtgtgctgggagaaatgt 240
QY 241 AATCATCTCTTCCACAACACTGTCATGTCCTGTGGGTGAACAGAAATCCCTGCCCT 300
DB 241 aatcatctcttccacaacactgtctcatgtccctgtggtggaacaaacaaatccctgccct 300
QY 301 CTCTGCCAGCAGACTGGTGGTCCAAAGATGGCAAAATGAGAGTGTGAAGGCTTC 360
DB 301 ctctgccagcagactgggtgggtccaaagatggcgaatgagagtgtgaaggcttc 360
QY 361 TTAGCGCAGTTGTTCAAGCCCTGGTGAATCTTAATCCAGTCCCTTACAAAGGCTAGA 420
DB 361 ttgagcgagttgttcagagccctgggtgatactgttaaccagtgccctaaaggctaga 420
QY 421 ACACTACAGGAGATGAATTTCTTAATAGAGCCGATGATCTGTGTCTTTGGACTCAT 480
DB 421 acaactacaggagatgaatttcttaataagagccgatgactgtgtctttgactcat 480
QY 481 CAAAGCCTTGTTAGCATTTGTCTGATTTTATCTTCAAAATTTCTCTGTGATTAAGAAGAT 540
DB 481 caaagccttgtagcatattgtcagtttattctcagaataattctctgtgattaagaagat 540

QY 541 AATTTATTAAAGTGATCCCTTCTTACTCTCTGTTGTTGTCGCCACAGACTTAGAAG 600
DB 541 aatttattaaagtgatcccttcttactctctggttgtgtgtgcgcacacagcttagaag 600
QY 601 TGCTATTAATAAAGAAAGAGCTCCAAATTTGAATACCTTATATTTACCATTTCTATAC 660
DB 601 tgctataataaagaaagagctcccaattgaatcacattataattaccattctataac 660
QY 661 AACAGCAGATGAGACGACTTTGAGACTTTTGCATGCTTATGTTGATCAGTAAAAA 720
DB 661 aacagcagatgagacgactttgagacttttgcattgcttattgattgacgtataaaaa 720
QY 721 GAATGTTACAGTAAACAATTAAGTGCAGTTTAA 754
DB 721 gaatgttacagttaacaataaagtgacgtttaaa 754

RESULT 2

AAAX87317
ID AAX87317 standard; cDNA; 754 BP.

XX AAX87317;

XX 27-SEP-1999 (first entry)

DE Human sensitive to apoptosis (SAG) gene mutant MM1.

XX SAG gene; sensitive to apoptosis; human; cancer; tumour;

KW neurodegenerative disease; muscular dystrophy; wound healing;

KW vulnerability; therapy; mutant; ds.

XX Homo sapiens.

OS Synthetic.

XX Key Location/Qualifiers

FT CDS 1..342

FT mutation /tag= a

FT replace(148,T)

FT /tag= b

FT /note= "C50S mutation"

PA (WARN) WARNER LAMBERT CO.

XX Sun Y;

XX WPI: 1999-430152/36.

DR P-PSDB: AAY06495.

PT SAG: Sensitive to Apoptosis Gene and related proteins, useful for
promoting cell growth and protecting cells against apoptosis

PS Claim 15; Page 59; 84pp; English.

XX This is the nucleotide sequence of human sensitive to apoptosis
CC (SAG) mutant gene MM1, which codes for a SAG protein (see AAY06495)
CC in which the Cys residue at position 50 of the native protein (see
CC AAY06492) is replaced by a Ser residue owing to a mutation of codon
CC 50 from TGC to AGC obtained by site-directed mutagenesis of SAG
CC cDNA. This residue is at a haem binding site of SAG. Single and
CC double SAG mutants (see AAX87317-31) were made in order to determine
CC the role of each cysteine residue of SAG in haem binding and SAG
CC oligomerization. MM1 showed reduced haem binding but unaffected
CC cell growth, protects cells from apoptosis, scavenges oxygen

[illegible]

RESULT	4
AXX87320	
ID	AXX87320 standard; cDNA; 754 BP.
XX	
AC	AXX87320;
XX	
DT	27-SEP-1999 (first entry)
XX	
DE	Human sensitive to apoptosis (SAG) gene mutant MMA:
KM	SAG gene: sensitive to apoptosis; human; cancer; tumour;
KW	neurodegenerative disease; muscular dystrophy; wound healing;
KM	vulnerable; therapy; mutant; ds.
XX	
OS	Homo sapiens.
OS	Synthetic.
XX	
FH	Location/Qualifiers
FT	1..342
FT	/*tag- a
FT	replace(181,T)
FT	mutation
FT	/*tag- b
FT	/note="C61S mutation"
XX	
PN	MO9932514-A2.
PN	
PD	01-JUL-1999.
XX	
XX	15-DEC-1998; 98WO-US26705
XX	
XX	

XX 11-SEP-1998; 98US-0099840.
PR 19-DEC-1997; 97US-0068179.
XX
PA (WARN) WARNER LAMBERT CO.
XX
PI Sun Y;
XX
XX WPI; 1999-430152/36.
DR P-PSDB; AAY06498.
DR
XX
XX SAG: Sensitive to Apoptosis Gene and related proteins, useful for
PT promoting cell growth and protecting cells against apoptosis
PT
XX
ES Claim 15, Page 64-65; 84pp; English.

CC This is the nucleotide sequence of human sensitive to apoptosis
CC (SAG) mutant gene MM4, which codes for a SAG protein (see AA066498)
CC in which the cys residue at position 61 of the native protein (see
CC AA066492) is replaced by a Ser residue owing to a mutation of codon
CC 50 from TGC to AGC obtained by site-directed mutagenesis of SAG
CC cDNA. This residue is at zinc-finger finger 1 of SAG. Single and
CC double SAG mutants (see AA087317-31) were made in order to determine
CC the role of each cysteine residue of SAG in haem binding and SAG
CC oligomerization. These properties were unaffected by the MM4
CC mutation. SAG is a novel zinc finger protein that promotes
CC cell growth, protects cells from apoptosis, scavenges oxygen
CC radicals and can be used for the reversal of a tumour phenotype.
CC SAG genes, and mutant SAG genes, can be used to protect cells from
CC apoptosis induced by redox reagents. They can also be used for the
CC recombinant production of SAG proteins, which are molecular targets
CC in the development of drugs against neurodegenerative disorders,
CC cancers and muscle dystrophy, and promoting wound healing.

Sequence 754 BP; 206 A; 155 C; 201 G; 192 T; 0 other;

Query Match	99.8%	Score 752.4	DB 20	Length 754
Best Local Similarity	99.98%	Pred. No. 7.3e-231		
Matches 753	Conservative	0	Mismatches 1	Indels 0
				Gaps 0

QY	1	ATATCCGACGCTGGAAAGACGGAGAGAAACCTGGGCCCTTGGCCCTCTCACTCCGGAGCTCA	60
Db	1	atggcgacgctlggaagacggagagaaacctgcgcccctggccctcactccggagactca	60
QY	61	GGCTCCAAAGTCGGGAGGCGCAAGAATTTCTTCCCTCAAGAAAGTGGAAAGCGGTGGGCATG	120
Db	61	ggtcccaagtctgggagggcgaagaagtcttcctcccaagaagtggaaacggtggtgcattg	120
QY	121	TGGAGCTGGGACGTGGAGTGGCATCTGCGCCATCTGCAGGGTCCAGGTGATGATGCC	180
Db	121	tggagcctgggacgtggaggtgcgatacgtgcacatctgcaggggtccaggtgcagtgatgctc	180
QY	181	TGTTTGAATGTCAAGCTGAAACAAACAAGAGACTGTTTGCTGGCTGGGAGAAATG	240
Db	181	agctctagatgctcaagctgaaacaaacaagaagactgctgtgtgtctctgggagaatgtc	240
QY	241	AATATTCTCTTCCACAACCTGCTGATGTCCCTTGGGTGAACAAGAACATCCCTGCCCT	300
Db	241	aatattctcttccacaacctgtgcattccctctgtgtgtgaacaagaacaatcgtcctccct	300
QY	301	CTCTGCCAGAGGACTGGGTGGTCCAAAGAATGGGCAAAAGACAGTGGTTAGAAGCGTTC	360
Db	301	ctctgccagagagactgggtgggtccaaagaatgggcaaaagacagtgggttagaaggcttc	360
QY	361	TTAGCGCAGTTGTTTCAGAGCCGCTGGTGGATCTTGTAATCCAGTGCCTTACAAAGGCTTAG	420
Db	361	tttagcgagttgtttcagagcccgctgggtggatcttgtaatccagtgcccttacaagaagctaga	420
QY	421	ACACTACAGGGGATGAATTCTTCAAAATAGAGCCGATGCATCTGTGCTCTTTGGACTCAT	480
Db	421	acctacaggggatgaaattcttcaaaatagagccgagtgatgcctcttggtcttcttggactcat	480

ID	AAx87324	strand; cDNA; 754 BP.
AC	AAx87324;	
DT	27-SEP-1999	(first entry)
DE	Human sensitive to apoptosis (SAG) gene mutant MM8.	
KW	SAG gene; sensitive to apoptosis; human; cancer; tumour; neurodegenerative disease; muscular dystrophy; wound healing; vulnerary; therapy; mutant; ds.	
OS	Homo sapiens.	
XX	Synthetic.	
FH	Key	Location/Qualifiers
FT	CDS	1..342
FT	mutation	/tag= a replace(262,T) /tag= b /note= "C885 mutation"
XX	WO9932514-A2.	
XX	01-JUL-1999.	
PD	15-DEC-1998;	98MO-US26705.
PF	11-SEP-1998;	98US-0099840.
PR	19-DEC-1997;	97US-0068179.
XX	(WARN) WARNER LAMBERT CO.	
PA	Sun Y;	
PI	WPT; 1999-430152/36.	
DR	P-PsDB; AAY06502.	
XX	SAG: Sensitive to Apoptosis Gene and related proteins, useful for promoting cell growth and protecting cells against apoptosis	
PT	Claim 15; Page 71; 84pp; English.	
PS	This is the nucleotide sequence of human sensitive to apoptosis (SAG) mutant gene MM8, which codes for a SAG protein (see AAY06502) in which the Cys residue at position 88 of the native protein (see AAV06492) is replaced by a Ser residue owing to a mutation of codon 88 from TGC to AGC obtained by site-directed mutagenesis of SAG cDNA. This residue (see AAx87317-31) were made in order to determine the role of each cysteine residue of SAG in haem binding and SAG oligomerization. These properties were unaffected by the MM8 mutation. SAG is a novel zinc finger protein that promotes cell growth, protects cells from apoptosis, scavenges oxygen radicals and can be used for the reversion of a tumour phenotype. SAG genes, and mutant SAG genes, can be used to protect cells from apoptosis induced by redox reagents. They can also be used for the recombinant production of SAG proteins, which are molecular targets in the development of drugs against neurodegenerative disorders, cancers and muscle dystrophy, and promoting wound healing.	
XX	Sequence 754 BP; 206 A; 155 G; 201 G; 192 T; 0 other;	

Query Match	99.8%;	Score 752.4;	DB 20;	Length 754;
Best Local Similarity	99.9%;	Pred. No. 7.3e-231;		
Matches 753; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;

0Y 1 ATGGCGCGAGCTGGGAAGACGGAGAGAAACCTGGCGCCTGGGCTCTCACTCGGAGAGCTCA 60
|||||
Db 1 atggcgcgacgtgggaagacggagagaaacctggcgccctgcctctaactccggagactca 60
|||||
0Y 61 GGCCTCAAGTCCGGGAGGCGACACATGTTCTCCCTTCAGAAAGTGGAAACGGGGTGGCCATG 120

[illegible]

RESULT	7
AAx87325	
ID	AAx87325 standard; cDNA; 754 BP.
XX	
AC	AAx87325;
XX	
DT	27-SEP-1999 (first entry)
XX	
DE	Human sensitive to apoptosis (SAG) gene
XX	
KW	SAG gene; sensitive to apoptosis; huma
KW	neurodegenerative disease; muscular dy
KW	vulnerable; therapy; mutant; ds.
XX	
OS	Homo sapiens.
OS	Synthetic.
XX	
FH	Key
FH	CDS
FT	Location/Qualifiers
FT	1..342
FT	/*tag= a
FT	replace(295,T)
FT	/*tag= b
FT	/note= "C9gs mutation
XX	
NN	WO9932514-A2.

XX 01-JUL-1999.
 PD 15-DEC-1998; 98WO-US26705.
 PF 11-SEP-1998; 98US-0099840.
 PR 19-DEC-1997; 97US-0068179.
 XX (WARN) WARNER LAMBERT CO.
 PA Sun Y;
 PI WPI; 1999-430152/36.
 DR P-PSDB; AAY06503.
 XX
 PT SAG: Sensitive to Apoptosis Gene and related proteins, useful for
 XX promoting cell growth and protecting cells against apoptosis
 XX
 PS Claim 15; Page 72-73; 84pp; English.
 XX
 CC This is the nucleotide sequence of human sensitive to apoptosis
 CC (SAG) mutant gene MM9, which codes for a SAG protein (see AAY06503)
 CC in which the Cys residue at position 99 of the native protein (see
 CC AAY06492) is replaced by a Ser residue owing to a mutation of codon
 CC 99 from TGC to AGC obtained by site-directed mutagenesis of SAG
 CC cDNA. This residue is in zinc-finger finger 2 of SAG. Single and
 CC double SAG mutants (see AAX87317-31) were made in order to determine
 CC the role of each cysteine residue of SAG in haem binding and SAG
 CC oligomerization. These properties were unaffected by the MM9
 CC mutation. SAG is a novel zinc finger protein that promotes
 CC cell growth, protects cells from apoptosis, scavenges oxygen
 CC radicals and can be used for the reversion of a tumour phenotype.
 CC SAG genes, and mutant SAG genes, can be used to protect cells from
 CC apoptosis induced by redox reagents. They can also be used for the
 CC recombinant production of SAG proteins, which are molecular targets
 CC in the development of drugs against neurodegenerative disorders,
 CC cancers and muscle dystrophy, and promoting wound healing.
 CC
 XX Sequence 754 BP; 206 A; 155 C; 201 G; 192 T; 0 other;

Query Match 99.8%; Score 752.4; DB 20; Length 754;
 Best Local Similarity 99.9%; Pred. No. 7.3e-231;
 Matches 753; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGCCGAGCTGGAGAGCGAAGAACTGCGCCCTGCTCTACTCCGGAGACTA 60
 DB 1 atggccgacgtggaagacggaagaaacctgcccctcctcctcaccgagactca 60
 QY 61 GGCTCCAGTGGGAGGAGCACAAGATGTTCCCTCAAGAGGAAAGCGGGGCCATG 120
 DB 61 ggctccaagtgggagggagacaagatgtctccctcaagaagtgaagcggggccatg 120
 QY 121 TGGAGCTGGAGCTGGAGTGCATACGTGCGCCATCTGAGGGTCCAGGTGATGATGCC 180
 DB 121 tggagctggagctggagtgcatatcgcccatctgaggggtccaggtgatgatgccc 180
 QY 121 tggagctggagctggagtgcatatcgcccatctgaggggtccaggtgatgatgccc 180
 DB 121 tggagctggagctggagtgcatatcgcccatctgaggggtccaggtgatgatgccc 180
 QY 181 TGTCTTAGATGTCAAGCTAAAAACAAGAGAGTGTGTGTGTGTGTGGGAGAAATGT 240
 DB 181 tgtcttagatgtcaagctaaaaacaagagagactgtgtgtgtgtgtgtgtgtgtgt 240
 QY 181 tgtcttagatgtcaagctaaaaacaagagagactgtgtgtgtgtgtgtgtgtgtgt 240
 DB 181 tgtcttagatgtcaagctaaaaacaagagagactgtgtgtgtgtgtgtgtgtgtgt 240
 QY 241 AATCATTCCTTCACAACTGCTGATGCTCCCTGGGGTGAACAGAACTGCTCCCT 300
 DB 241 aatcatctcttcacaaactgctgcatgtccctgggggtgaacagaaactgctccct 300
 QY 301 CTCTGCGACGAGACTGGGTGTCCTCAAGATGCGCAAAATGAGAGTGTAGAGGCTTC 360
 DB 301 ctctgcgacgagactgggtgtctcctcaagatgcgcaaaatgagagtgttagagcttc 360
 QY 361 TTAGCGCAGTGTGTCGAGCCCTGGGAGATCTTGTAACTCAGCCCTCAAGAGCTA 420
 DB 361 tttagcgactgtgtcagagccctgggagatcttgttaactcagccctcaagagctaga 420

QY 421 ACACTACAGGGAGTGAATTCCTCAATAGAGCCGATGATCTGTGCTTTGGACTCAT 480
 DB 421 acactacagggagtgatgaattcctcaatagagccgatgattctgtgtcttggactcat 480
 QY 481 CAAAGCCTTGTTAGCATTGTTCAGTTTATCTTCAGAAATTTCTGTGATTAAGAGAT 540
 DB 481 caaagccttggttagcattgttccagtttattcttcagaaatttctgtgattaaagat 540
 QY 541 AATTTTAAAGTGCTGCTTCTTACCTGAGTGCTGTGCTGGCAGACAGCTTTGAAG 600
 DB 541 aattttaaagtgtgcttcttacctgagtgctgtgctggcagacagctttgaag 600
 QY 601 TCGTATATAAAGGAAGAGCTCAATTAATGAATCACCCTATATATTAATTAATTAAC 660
 DB 601 tgcataataaaggaaagagccccaatggaatcaccctataatattaccattctatcc 660
 QY 661 AACAGGCACTGGAAGCAGTTTCGAGACTTTTCGATGCTTATGATGATCACTTAAAAA 720
 DB 661 aacaggcactggaaagcagtttcgagacttttcgattgattatgattgattgattgatt 720
 QY 721 GAATGTTACAGTACAAATTAAGTGCAGTTTAA 754
 DB 721 gaatgttacagtacaaatgaagtgcagttttaa 754
 RESULT 8
 AAX87327
 ID AAX87327 standard; cDNA; 754 BP.
 AC AAX87327;
 XX
 XX 27-SEP-1999 (first entry)
 DE Human sensitive to apoptosis (SAG) gene mutant MM1.
 XX
 XX SAG gene; sensitive to apoptosis; human; cancer; tumour;
 KW neurodegenerative disease; muscular dystrophy; wound healing;
 KW vulnerability; therapy; mutant; ds.
 XX
 OS Homo sapiens.
 OS Synthetic.
 OS
 FH Key Location/Qualifiers
 FT CDS 1..342
 FT mutation /*tag= a
 FT /*tag= b
 FT /*note= "C64s mutation"
 PN MO9932514-A2.
 PD 01-JUL-1999.
 XX
 XX 15-DEC-1998; 98WO-US26705.
 PF 11-SEP-1998; 98US-0099840.
 PR 19-DEC-1997; 97US-0068179.
 XX
 PA (WARN) WARNER LAMBERT CO.
 PA Sun Y;
 PI WPI; 1999-430152/36.
 DR P-PSDB; AAY06503.
 XX
 PT SAG: Sensitive to Apoptosis Gene and related proteins, useful for
 PT promoting cell growth and protecting cells against apoptosis
 XX
 PS Claim 15; Page 76; 84pp; English.
 XX
 CC This is the nucleotide sequence of human sensitive to apoptosis
 CC (SAG) mutant gene MM1, which codes for a SAG protein (see AAY06503)
 CC in which the Cys residue at position 64 of the native protein (see

CC AAY06492) is replaced by a Ser residue owing to a mutation of codon
 CC 64 from TGT to AGC obtained by site-directed mutagenesis of SAG
 CC cDNA. This residue is in zinc-finger finger 1 of SAG. Single and
 CC double SAG mutants (see AAX87317-31) were made in order to determine
 CC the role of each cysteine residue of SAG in haem binding and SAG
 CC oligomerization. These properties were unaffected by the MM12
 CC mutation. SAG is a novel zinc finger protein that promotes
 CC cell growth, protects cells from apoptosis, scavenges oxygen
 CC radicals and can be used for the reversion of a tumour phenotype.
 CC SAG genes, and mutant SAG genes, can be used to protect cells from
 CC apoptosis induced by redox reagents. They can also be used for the
 CC recombinant production of SAG proteins, which are molecular targets
 CC in the development of drugs against neurodegenerative disorders,
 CC cancers and muscle dystrophy, and promoting wound healing.
 CC
 XX
 XX
 S0 Sequence 754 BP; 206 A; 155 C; 201 G; 192 T; 0 other;

Query Match 99.8%; Score 752.4; DB 20; Length 754;
 Best Local Similarity 99.9%; Pred. No. 7.3e-231;
 Matches 753; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGCGCCAGCTGGAAGAGGAGAGAAACCTGGCCCTGCTCTCACTCCGGAGCTCA 60
 Db 1 atgcccagctggaagagagagagaaacccgcccctgacctccactccggagctca 60
 QY 61 GGTCTCAAGTGGAGGCGACAGATGTTCTCCCTCAAGAGTGAACGGGTCGCATG 120
 Db 61 ggtcccaagtcggagggagagaaagatgttccctcaagaagtgaacggcggtgcattg 120
 QY 121 TGGAGCTGGGAGCTGGAGTGGATACGTGGCCATCTGCAGGGTCCAGGTGATGATCC 180
 Db 121 tggagctgggagctggagtggtgagacgtgcacatccgaggtccagtgatgagcc 180
 QY 181 TGTCTTAGATCTCAAGCTGAAACAAACAAGAGACTGTGTGTGGTGGGGAGATGT 240
 Db 181 tgtcttagaatgtaagctgaaacaaacaagagactgtgtgtgtggtgggagatgt 240
 QY 241 AATCATCTCTTCCACAACTGCTGCATGCTCCCTGTGGTGAACAGAACTCGTCCCT 300
 Db 241 aatcattctctcacaaactgtcatgtccctgtggtgagaaacagaaactcgtgccct 300
 QY 301 CTCTGCCAGAGAGCTGGGTGGTCCAAAGAAATGGCAAAATGAGAGTGTGATGAAGCTTC 360
 Db 301 ctctgccagagagctgggtgggtccaaagaaatcgcaaatgagatggtgaagagcttc 360
 QY 361 TTAGCGCAGTGTGTACAGAGCCCTGTGGATCTGTAATCCAGTCCCTACAAAGGCTAGA 420
 Db 361 tttagcgagtggtgtacagagccctgtggatctgtaatccagtccttaaaagctaga 420
 QY 421 ACACATACAGGGGATGAATTTTCAATAGAGCCGATGATGATGTGTGTTGGACTCAT 480
 Db 421 acctatacagggatgaattttcaatagagccgatgattgtgtgttcttgactcat 480
 QY 481 CAAGCGCTGGTGAAGCTTGTGAGTTTATCTTCAGAAATTTCTCTGTGATTAAGAAGAT 540
 Db 481 caaagcctggtagaagcttgtgagtttatacttcagaaattctcgtgattaagaagat 540
 QY 541 AATTATTAAGGTGTCTTCTTACCTCTGTTGGTGTGTGTGTCGACACAGCTTGAAG 600
 Db 541 aattattaaagtggtcttcttacctctcgtgtgtgtgtgtgctgcgacacagcttgaag 600
 QY 601 TGTGTAATAAAGGAGAAAGAGCTCCAAATTAATGAAATCACCATTATAATTAACCATTTATAC 660
 Db 601 tgtataaaaaaagaaagagctccaattgaatcacctataattaccattctatc 660
 QY 661 AACAGCGATGGAAGAGAGCTTGTGAGACTTTTGTGAGCTTATGATGATCAGTTAAAAA 720
 Db 661 aacagcgatggaaagagagcttgtgagacttttgcagcttatgtgatcagcttaaaaaa 720
 QY 721 GAATGTACAGTAACAATAAATTAAGTGCAGTTTAAA 754
 Db 721 gaatgtacagtaacaataaataaagtgcagtttaa 754

RESULT 9
 ID AAX87328
 XX AAX87328 standard; cDNA; 754 BP.
 AC AAX87328;
 XX
 DT 27-SEP-1999 (first entry)
 DE Human sensitive to apoptosis (SAG) gene mutant MM12.
 XX
 KW SAG gene; sensitive to apoptosis; human; cancer; tumour;
 KW neurodegenerative disease; muscular dystrophy; wound healing;
 KM veterinary; therapy; mutant; ds.
 XX
 OS Homo sapiens.
 OS Synthetic.
 FH Key Location/Qualifiers
 FT CDS 1..342
 FT mutation /tag= a
 FT /tag= b
 FT /note= "C73S mutation"
 PN MO9932514-A2.
 PD 01-JUL-1999.
 XX
 PF 15-DEC-1998; 98MO-US26705.
 XX
 PR 11-SEP-1998; 98US-0099840.
 PR 19-DEC-1997; 97US-0068179.
 XX
 PA (WARN) WARNER LAMBERT CO.
 XX
 PI Sun Y;
 XX
 DR WPI; 1999-430152/36.
 DR P-PSDB; AAY06506.
 XX
 PS Claim 15; Page 77-78; 84pp; English.
 CC This is the nucleotide sequence of human sensitive to apoptosis
 CC (SAG) mutant gene MM12, which codes for a SAG protein (see AAY06506)
 CC in which the Cys residue at position 73 of the native protein (see
 CC AAY06492) is replaced by a Ser residue owing to a mutation of codon
 CC 73 from TGT to AGC obtained by site-directed mutagenesis of SAG
 CC cDNA. This residue is in a protease inhibitor motif of SAG. Single
 CC and double SAG mutants (see AAX87317-31) were made to determine
 CC the role of each cysteine residue of SAG in haem binding and SAG
 CC oligomerization. These properties were unaffected by the MM12
 CC mutation. SAG is a novel zinc finger protein that promotes
 CC cell growth, protects cells from apoptosis, scavenges oxygen
 CC radicals and can be used for the reversion of a tumour phenotype.
 CC SAG genes, and mutant SAG genes, can be used to protect cells from
 CC apoptosis induced by redox reagents. They can also be used for the
 CC recombinant production of SAG proteins, which are molecular targets
 CC in the development of drugs against neurodegenerative disorders,
 CC cancers and muscle dystrophy, and promoting wound healing.
 CC
 XX
 S0 Sequence 754 BP; 206 A; 155 C; 201 G; 192 T; 0 other;

Query Match 99.8%; Score 752.4; DB 20; Length 754;
 Best Local Similarity 99.9%; Pred. No. 7.3e-231;
 Matches 753; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGCGCCAGCTGGAAGAGGAGAGAAACCTGGCCCTGCTCTCACTCCGGAGCTCA 60

```

Db      1 atggccgaagctgtaagaacgagagaacacgtgcgccttcctcaccgcggagctca 60
OY      61 GGCTCCAAATCGGGAGGCGACAAGATGTTCTCCCTCAGAAGTGCAGCGGGTGCCTATG 120
Db      61 ggctccaagtcgagggcgacaagaatgcttcctcctaagaatggaacgcgltgacatg 120
OY      121 TGGAGCTGGAGCTGGAGTGCATACGTGCGCATCTCGAGGCTCCAGGTGATGATGTCG 180
Db      121 tggagctggagctggagtgagtgacacgtagcgccatctcgaggtccaggtgtagtgc 180
OY      181 TGTCTTAGATGTCAGAGCTGAAAAAACAAGAGAGACTGTGTGTGCTGGGAGATGT 240
Db      181 tgtcttagatgtaagctgtaaaacaagaagagacagtggtgtgtcgggagagatgt 240
OY      241 AATCATTCCTTCACAACTGCTGCATGTCCTGCTGGGTGAACAGAACAAATGCTCCCT 300
Db      241 aatcatctcttcacaaactgctgcatgtccctgtgtgtaaacagaaacatcgctgcct 300
OY      301 CTCTGCCAGCAGACTGGGTGGTCCAAAGATCGGCAATGAGAGTGTAGAGGCTTC 360
Db      301 ctctgccagcagactgggtgtgtccaaagaatcgccaatgagatggttagaaggcttc 360
OY      361 TTAGCGCAGTGTGTCAGAGCCCTGTGTGATCTTGAATCAGTGCCTCAAAAGGCTAGA 420
Db      361 ttagcgcaagtgttcagagccctgtgtgacttgaatccagtgccctacaaggctaga 420
OY      421 AACACACAGGGGATGAATCTTCAAAATAGAGCCGATGATGTGTGCTTGGACTCAT 480
Db      421 aaacacaggggagatgaatcttcaaaatagagccgatgatatgtgttcttgactcat 480
OY      481 CAAGACCTTGTTACATTTGATTTTATCTTCAAGAAATCTCTGATTAAGAGAT 540
Db      481 caagaccttggttacatttgcattgtccaagttatcttcagaatctctgtgataaggat 540
OY      541 AATTATTAAGGTGTCCTTCTCACTCTGTGTGTGTGTGCGCACACACAGCTTAGAAG 600
Db      541 aattatataaggtgtctctcctcctcctctgtgtgtgtgtgcgcacacagcttagaag 600
OY      601 TGCTTTAAAAAGGAAGGCTCAAAATGATCACTTATTAATTAATTAATTAATTAATTA 660
Db      601 tgccttataaaaggaagagctccaatgaaatgaaatcctataatcctcctctatc 660
OY      661 AACAGGACAGTGGAGACAGTTTGCAGACTTTTGCATGTTAGTTGATCAGTTAAAAA 720
Db      661 aacaggacagtggaagacagtttcgagacttttcgatatgtatgattcagttaaaaa 720
OY      721 GAATGTTACAGTACCAATTAAGTGCAGTTTAAA 754
Db      721 gaatgttacagtaacaataaagtcagttaaa 754

```

RESULT 10

AAK87331
ID AAK87331 standard: cDNA: 754 BP.

AC AAK87331:

DT 27-SEP-1999 (first entry)

XX Human sensitive to apoptosis (SAG) gene mutant MM15.

XX SAG gene: sensitive to apoptosis; human: cancer; tumour;

KW neurodegenerative disease: muscular dystrophy; wound healing;

XX vulnerability; therapy; mutant; ds.

OS Homo sapiens.

XX Synthetic.

XX Key

FT CDS

FT mutation

Location/Qualifiers

1..342
/*tag-
replace(139,C)

```

FT      /*tag= b
FT      /note= "C47S mutation"
PN      W09332514-A2.
PD      01-JUL-1999.
PF      15-DEC-1998; 98WO-US26705.
PR      11-SEP-1998; 98US-0099840.
PR      19-DEC-1997; 97US-0068179.
PA      (WARN ) WARNER LAMBERT CO.
PI      Sun Y;
DR      WPI: 1999-430152/36.
DR      P-PSDB; AAY06509.
PT      SAG: Sensitive to Apoptosis Gene and related proteins, useful for
PS      promoting cell growth and protecting cells against apoptosis
PS      Claim 15; Page 82-83; 84pp; English.

This is the nucleotide sequence of human sensitive to apoptosis
(SAG) mutant gene MM15, which codes for a SAG protein (see AAY06509)
in which the Cys residue at position 47 of the native protein (see
AAY06492) is replaced by a Ser residue owing to a mutation of codon
47 from TGC to AGC obtained by site-directed mutagenesis of SAG
cDNA. This residue is in a GMPH binding site of SAG. Single
CC and double SAG mutants (see AAK87317-31) were made to determine
the role of each cysteine residue of SAG in haem binding and SAG
oligomerization. These properties were unaffected by the MM15
mutation. SAG is a novel zinc finger protein that promotes
cell growth, protects cells from apoptosis, scavenges oxygen
radicals and can be used for the reversal of a tumour phenotype.
CC SAG genes, and mutant SAG genes, can be used to protect cells from
apoptosis induced by redox reagents. They can also be used for the
recombinant production of SAG proteins, which are molecular targets
in the development of drugs against neurodegenerative disorders,
cancers and muscle dystrophy, and promoting wound healing.

Sequence 754 BP; 206 A; 155 C; 201 G; 192 T; 0 other;

Query Match      99.8%; Score 752.4; DB 20; Length 754;
Best Local Similarity 99.9%; Pred. No. 7.3e-231;
Matches 753; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```


OY 661 AACAGCAGTGGAGACGTTTGCAGCTTTTGCATTGATGATCACTTAAAAA 720
 |||
 DB 661 aacagcagtggaagcagcttcgcagcttcgcagcttcgcttaagctcagcttaaaaa 720
 OY 721 GAATGTTACAGTACCAATAAAGTCAGTTTAA 754
 |||
 DB 721 gaatgtacagtaacaataaagtcagtttaa 754

 RESULT 12
 AAX87322
 ID AAX87322 standard; cDNA; 754 BP.
 XX
 AC AAX87322;
 XX
 DT 27-SEP-1999 (first entry)
 XX
 DE Human sensitive to apoptosis (SAG) gene mutant MM6.
 XX
 KW SAG gene; sensitive to apoptosis; human; cancer; tumour;
 KW neurodegenerative disease; muscular dystrophy; wound healing;
 KW vulnerability; therapy; mutant; ds.
 OS Homo sapiens.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..342
 FT /*tag= a
 FT mutation replace(244..246,CAT)
 FT /*tag= b
 FT /note= "H82K mutation"
 XX
 PN WO932514-A2.
 XX
 PD 01-JUL-1999.
 XX
 PE 15-DEC-1998; 98WO-US26705.
 XX
 PR 11-SEP-1998; 98US-0099840.
 PR 19-DEC-1997; 97US-0068179.
 XX
 PA (WARN) WARNER LAMBERT CO.
 XX
 PI Sun Y;
 DR WPI; 1999-430152/36.
 DR P-PSDB; AAY06500.
 XX
 PT SAG: Sensitive to Apoptosis Gene and related proteins, useful for
 PT promoting cell growth and protecting cells against apoptosis
 XX
 PS Claim 15; Page 67-68; 84pp; English.
 XX
 CC This is the nucleotide sequence of human sensitive to apoptosis
 CC (SAG) mutant gene MM6, which codes for a SAG protein (see AAY06500)
 CC in which the His residue at position 82 of the native protein (see
 CC AAY06452) is replaced by a Lys residue owing to a mutation of codon
 CC 82 from CAT to AAA obtained by site-directed mutagenesis of SAG
 CC cDNA. This residue is in zinc-finger 1 of SAG. Single and
 CC double SAG mutants (see AAX87317-31) were made in order to determine
 CC the role of each cysteine residue of SAG in haem binding and SAG
 CC oligomerization. These properties were unaffected by the MM6
 CC mutation. SAG is a novel zinc finger protein that promotes
 CC cell growth, protects cells from apoptosis, scavenges oxygen
 CC radicals and can be used for the reversion of a tumour phenotype.
 CC SAG genes, and mutant SAG genes, can be used to protect cells from
 CC apoptosis induced by redox reagents. They can also be used for the
 CC recombinant production of SAG proteins, which are molecular targets
 CC in the development of drugs against neurodegenerative disorders,
 CC cancers and muscle dystrophy, and promoting wound healing.
 XX

SQ Sequence 754 BP; 207 A; 154 C; 201 G; 192 T; 0 other;

 Query Match 99.6%; Score 750.8; DB 20; Length 754;
 Best Local Similarity 99.7%; Pred. No. 2.4e-230;
 Matches 752; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

 OY 1 ATGGCGACGTGGAAGACGAGGAAACCTCGCCGCTGCTCAGTCCGGAGCTCA 60
 |||
 DB 1 atggcgcagctggaagacggaagaaacccgcgcctccatccgcggagctca 60
 OY 61 GGCTCCAACTCGGAGGCGACAAAGATGTTCTCCCTCAAGAACTGGAACGCGTGCAT 120
 |||
 DB 61 ggctccaaactcggagcgacaaagatgtctccccaagaagtggaacgcggtgcctatg 120
 OY 121 TGGAGCTGGAGACGTGAGTGCATACGTGCGCATCTGAGAGGTCCAGGTATGATGCC 180
 |||
 DB 121 tggagctggagacgtgagtgatgcatacgtgcacatcgaagtgatgagtgcc 180
 OY 181 TGTCTTAGATGTCAGAGCTGAACAAACAAAGAGAGACTGTGTGTGGGAGAAATGT 240
 |||
 DB 181 tgtcttagatgtcagagctgaacaaacaaagagagactgtgtgtcgggagaaagt 240
 OY 241 AATCATTCCTTCACAACTGCTGATGCTCCGTGGGTGAACAGAAATGCTGCCCT 300
 |||
 DB 241 aataatctctccacaactgctgctccctgtgggtgaacagaaatcgcctgcct 300
 OY 301 CTCTGCACAGACGACTGGGTGCTCCAAAGATCGCAATGAGTGGTTAGAGGCTTC 360
 |||
 DB 301 ctctgcacagactgggtgctccaaagatcgcgaatgagtggttagaagcttc 360
 OY 361 TTAGCGCAGTGTTCAGAGCCCTGCTGATCTTGAATCCATGCTCAAGAGCTAGA 420
 |||
 DB 361 ttagcgagtggttcagagccctgctgatatcttgaaatccatgctcacaagctaga 420
 OY 421 ACAGTACAGGGAGTGAATCTTCAATATAGAGCCAGTGAATGTGTGCTTGGACTCAT 480
 |||
 DB 421 acactacagggagtgatctctcaaaatagagcgaatgactgtgtcttgactcat 480
 OY 481 CAAAGCCTGTTAGCATTTGTCAGTTTATCTCAGAAATCTCTGATTAAGAGAT 540
 |||
 DB 481 caaagcctgttagcatcttgcagtttatacttcaagaatctctgtatcaagaagat 540
 OY 541 AATTATTAAAGTGCTCTCTCTACCTCTGCTGATGCTGCGCACACAGCTTAGAG 600
 |||
 DB 541 aattattaaagtgtctctctctactctctgtgtgtgtgtgcgcacacagcttagag 600
 OY 601 TCGTATATAAAGAAAGAGCTCCAAATGAATACCTTAATATTTACCAATTTCTATAC 660
 |||
 DB 601 tgcataaaaaagaaagagctcccaaatgaaatcactataattaccattctatag 660
 OY 661 AACAGCAGTGGAGACGTTTGCAGACTTTTGCAGTGTGATGATCACTTAAAAA 720
 |||
 DB 661 aacagcagtggaagcagcttcgcagcttcgcagcttcgcttaagctcagcttaaaaa 720
 OY 721 GAATGTTACAGTACCAATAAAGTCAGTTTAA 754
 |||
 DB 721 gaatgtacagtaacaataaagtcagtttaa 754

 RESULT 13
 AAX87323
 ID AAX87323 standard; cDNA; 754 BP.
 XX
 AC AAX87323;
 XX
 DT 27-SEP-1999 (first entry)
 XX
 DE Human sensitive to apoptosis (SAG) gene mutant MM7.
 XX
 KW SAG gene; sensitive to apoptosis; human; cancer; tumour;
 KW neurodegenerative disease; muscular dystrophy; wound healing;
 KW vulnerability; therapy; mutant; ds.
 KW

XX OS Homo sapiens.
 OS Synthetic.
 XX FH Key
 FT CDS Location/Qualifiers
 FT mutation 1..342
 FT /tag= a
 FT replace(253..255,CAC)
 FT /tag= b
 FT /note= "H85K mutation"
 XX PN W09932514-A2.

XX PD 01-JUL-1999.
 XX PF 15-DEC-1998; 98WO-US26705.
 XX PR 11-SEP-1998; 98US-0099840.
 XX PR 19-DEC-1997; 97US-0068179.
 XX PA (WARN) WARNER LAMBERT CO.
 XX PI Sun Y;
 XX DR WPI: 1999-430152/36.
 XX DR P-PSDB; AAY06501.

PT SAG: Sensitive to Apoptosis Gene and related proteins, useful for
 PT promoting cell growth and protecting cells against apoptosis
 PS Claim 15; Page 69-70; 84pp; English.

CC This is the nucleotide sequence of human sensitive to apoptosis
 CC (SAG) mutant gene MM7, which codes for a SAG protein (see AAY06501)
 CC in which the His residue at position 85 of the native protein (see
 CC AAY06492) is replaced by a Lys residue owing to a mutation of codon
 CC 85 from CAC to AAG obtained by site-directed mutagenesis of SAG
 CC cDNA. This residue is in zinc-finger finger 2 of SAG. Single and
 CC double SAG mutants (see AAX87317-31) were made in order to determine
 CC the role of each cysteine residue of SAG in haem binding and SAG
 CC oligomerization. These properties were unaffected by the MM7
 CC mutation. SAG is a novel zinc finger protein that promotes
 CC cell growth, protects cells from apoptosis, scavenges oxygen
 CC radicals and can be used for the reversion of a tumour phenotype.
 CC SAG genes, and mutant SAG genes, can be used to protect cells from
 CC apoptosis induced by redox reagents. They can also be used for the
 CC recombinant production of SAG proteins, which are molecular targets
 CC in the development of drugs against neurodegenerative disorders,
 CC cancers and muscle dystrophy, and promoting wound healing.

XX Sequence 754 BP; 206 A; 153 C; 202 G; 193 T; 0 other;

Query Match 99.6%; Score 750.8; DB 20; Length 754;
 Best local similarity 99.7%; Pred. No. 2..4e-230;
 Matches 752; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGCCGACGTGGAAGAAGCGAGAGAAACCTGGCCCTCTCTACTCCGGAGCTCA 60
 DB 1 atggccgacgtggaagaagcgagagaaacctggccctctctactccggagctca 60
 QY 61 GGCTCCCAAGTCGGAGCGACGACAGATGTTCTCCCTCAAGAAGTGAAGCGGGGCGCAG 120
 DB 61 ggctcccaagtcggagcgacgacgacagatgttctccctcaagaagtgaagcgggcgccag 120
 QY 121 TGGAGCTGGGACGTGAGTGCATACGTGCGCCATCTGCAGGCTCCAGTGATGATGCC 180
 DB 121 tggagctgggacgtgagtgcatatcgctgcgcacatcgagggctccagtgatgatgcc 180
 QY 181 TGGCTAGAGTGTCAAGCTGAAAACAACAAGAGACTGTGTGTGCTGGGAGATGT 240
 DB 181 tggcttagagtgtcaagctgaaaacaacaagagactgtgtgtgctgggagatgt 240

QY 241 AATCATCTCTTCCACCAACTGCTGCATGTCCCTGTGGGTGAAGAAACAATCGTCGCT 300
 DB 241 aatcattctcttccaccaaactgctgcattgtccctgtgggtgaagaaacaatcgctgcct 300
 QY 301 CTCTGCCAGCAGGACCTGGGTGTCACAAAGAAATGGCAAAATGAGATGCTTAGAAGCTTC 360
 DB 301 ctctgccagcagacctgggtgtctacaaagaaatggcaaaatgagatgcttagaagcttc 360
 QY 361 TTACGCGAGTGTTCAGAGCCCTGTGGATCTTGAATCCAGTGCCTCAAAAGCTTAGA 420
 DB 361 ttacgcgagtggttcagagccctgtggatcttgaatccagtgcctcaaaagcttaga 420
 QY 421 AACACTACAGGGATGATTTCTTCAATAGAGCCGATGATCTGTGCTCTTTGGACTCAT 480
 DB 421 aacactacagggatgatttcttcaatagagccgatgatctgtgctctttggactcat 480
 QY 481 CAAAGCCTTGTTAGCATTTGTGCTGTTATCTTCAGAAATTCCTGTGATTAAGAAGAT 540
 DB 481 caaagccttgttagcatttgtgctgttattcttcagaaattcctgtgattagaagat 540
 QY 541 AATTTATTAAGGTGCTCTCTCTACCTGTGTGTGTGTCGCGACACACCTTAGAAG 600
 DB 541 aatttatttaagtgctctctctacacctgtgtgtgtgtgtcgcgacacaccttagaag 600
 QY 601 TGCATATAAAGAGAAAGAGCTCCAAATGAATGAACTTATATTAATTAACCATTTCTATAC 660
 DB 601 tgcataataaagagaaagagctccaaatgaatgaacttattatattaatcatcttatac 660
 QY 661 AACAGCAGTGGAGAGCAGTTTCGAGACTTTTCGATGCTTATGTTGATCAGTTAAAAA 720
 DB 661 aacagcagtggaagagcagtttcgagacttttcgatgcttattggttgcagttaaaaa 720
 QY 721 GAATGTTACAGTACCAATTAAGTGCAGTTTAA 754
 DB 721 gaatgttacagtaccaatthaagtgcagttttaa 754

RESULT 14

AAH87329
 ID AAX87329 standard; cDNA; 754 BP.

XX AAX87329;

AC 27-SEP-1999 (first entry)

XX Human sensitive to apoptosis (SAG) gene mutant MM13.

XX SAG gene; sensitive to apoptosis; human; cancer; tumour;

XX neurodegenerative disease; muscular dystrophy; wound healing;

XX vulnerable; therapy; mutant; ds.

XX Homo sapiens.

OS Synthetic.
 OS Synthetic.
 XX Key
 XX CDS Location/Qualifiers

FT mutation 1..342
 FT /tag= a
 FT replace(181,T)
 FT /tag= b
 FT /note= "C61S mutation"
 FT replace(190,T)
 FT /tag= c
 FT /note= "C64S mutation"

PN W09932514-A2.

XX PD 01-JUL-1999.

XX PF 15-DEC-1998; 98WO-US26705.

XX PR 11-SEP-1998; 98US-0099840.

XX PR 19-DEC-1997; 97US-0068179.

PA (WARN) WARNER LAMBERT CO.

PI Sun Y;

DR WPI; 1999-430152/36.

DR P-PSDB; AAY06507.

PT SMG: Sensitive to Apoptosis gene and related proteins, useful for promoting cell growth and protecting cells against apoptosis

Claim 15; Page 79-80; 84pp; English.

This is the nucleotide sequence of human sensitive to apoptosis (SAG) mutant gene MM3, which codes for a SAG protein (see AY06507). In which the Cys residues at positions 61 and 64 of the native protein (see AY06492) are replaced by Ser residues owing to mutations of codons 61 and 64 through site-directed mutagenesis of SAG cDNA. These residues are in zinc ring finger 1 of SAG. Single double SAG mutants (see AA87317-31) were made in order to determine the role of each cysteine residue of SAG in haem binding and SAG oligomerization. MM3 showed greatly reduced haem binding (and a decreased brown colour) but oligomerization was unaffected. SAG is a novel zinc finger protein that promotes cell growth, protects cells from apoptosis, scavenges oxygen radicals and can be used for the reversion of a tumour phenotype. SAG genes, and mutant SAG genes, can be used to protect cells from apoptosis induced by redox reagents. They can also be used for the recombinant production of SAG proteins, which are molecular targets in the development of drugs against neurodegenerative disorders, cancers and muscle dystrophy, and promoting wound healing.

SQ Sequence 754 BP; 207 A; 155 C; 201 G; 191 T; 0 other;

Query Match	Score	DB	Length
99.68;	750.8;	20;	754;

Matches	752;	Conservative	0;	Mismatches	2;	Indels	0;	Gaps	0;
---------	------	--------------	----	------------	----	--------	----	------	----

OY	1	ATGCCAGCTGGGAAGCGGAGAGGAAACCTGGCCCTGGCCCTCAATCCGGAGACTGA	60
Db	1	atggccagcttgaagagcggagagaagaaccccgccctggcctctcaaccgcggagacttca	60
OY	61	GGCTCCAAAGTCGGGAGGCGACGAAGAATGTTCTCCCTCAAGAAGTGAACCGGCTGGCCATG	120
Db	61	ggctccaagtcggagagcgcgacaagatgcttccctccaagaatgagaaacgcggtggccatg	120
OY	121	TGGAGCTGGGACGTCGAGCTCGATACGTGGCCATCTGCAGAGGTCCAGGTTGATGATGCC	180
Db	121	tggagctgggacgtcgagctcgcgatcgtggccatctgcagaggtccaggttgatgattgcc	180
OY	181	TGCTTTAGATGTCAAAGCTGMAAAACAACAAGAGACATGTTGTGTGGTCTGGGGAGAAATG	240
Db	181	agctctgaagtcacgctgaagaaacaacaagaagactgctgtgtgtctcgggagaaatgt	240
OY	241	AATCATTCCTTCACACACTCTCTCATGTCCCTCTGGGTGAACAGAAACAATGCTGCTCCT	300
Db	241	aatcatctccttcacacactctctcatgtccctctgggtgaacagaaacaatatgctgctcct	300
OY	301	CTCTGCACGACGACTGGGTGTCTCCAAGAAATGGCAAAATGAGAGTGTTAGAAGGCTTC	360
Db	301	ctctgcacgacgactgggtgtctccaagaaatggcaaaatgagagtgttagaaggcttc	360
OY	361	TTAGACGAGTTGTTACAGAGCCCTGGGTGATCTTTGTAATCAGTGCCTTCAAAAGCTAGA	420
Db	361	ttagacgagttgttacagagccctgggtgatctttgtaatcagtgccttcaaaagctaga	420
OY	421	ACAATACAGGCGATGTAATTTCTCAAAATAGAGCGCATGTGATCTGTGATTTGGACTCAT	480
Db	421	acaatacaggcgatgtaatcttccaataagagcgatgatactgtgtcttggactcat	480
OY	481	CAAAAGCTTGATAGCAATTTGTCAAGTTTATCTTCAAGAAATTTCTCTGATTTAAGAAAT	540
Db	481	caaaagcttggatagcaatttgtcaagtttatcttcaagaaatctctctgatttaagaagat	540

QY	541	AAATTTAAAGGAGGTCCTTCACCTCTGTGTGTGTCGGGACACAGGTTAAAG	6000
QY	541	AA	
Db	541	aattatttaaagaagtgctctctccacccttggtgctgctgcgcacacagcttaaga	6000
QY	601	TGCTATTAAGGAGGAGGCTCCCAATTGNAATACCTTAATTAATACCATTTCTTAC	6600
Db	601	tgcataaaaaagaagaagtgctccaaattgatacctaataattacccaattctatac	6600
QY	661	AACAGCAGTGGAAAGCAGTTTCGAGACTTTTTCGATGCTTAATGTTATCAGTTAAAAA	7200
Db	661	aacagcgagtcggaagcagtttcgagaccttttcgattgcttattgltatcagttaaaaaa	7200
QY	721	GAATGTTACAGTAACAAATAAAGGACAGTTTAAA	754
Db	721	gaatgltacagtaaacaaataaagtcgagttttaa	754

RESULT 15

ID AAX87330 standard; cDNA; 754 BP.

AC AAX87330;

DT 27-SEP-1999 (first entry)

Human sensitive to apoptosis (SAG) gene mutant MML4.

KW SAG gene; sensitive to apoptosis; human; cancer; tumour;

KW vulnerable; therapy; mutant; ds.

OS Homo sapiens.

XX
XX

FT	CDS
----	-----

FT mutation

LE

ET

XX
000000

XX

XX
XX

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

PR 19-DEC-1997;

PA (WARN) WARNH

PI Sun Y;

DR WPI; 1999-430

XX

PT promoting cell

PS Claim 15; Page 22

CC This is the nucleotide sequence of human sensitive to apoptosis
CC (SAG) mutant gene MM14, which codes for a SAG protein (see AY06508).
CC in which the Cys residues at positions 99 and 107 of the native
CC protein (see AY06492) are replaced by Ser residues owing to
CC mutations of codons 99 and 102 through site-directed mutagenesis of
CC SAG cDNA. These residues are in zinc ring finger 2 of SAG. Single
CC double SAG mutants (see AA87317-13) were made in order to determine

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 25, 2002, 14:08:41; Search time 3583.2 Seconds

(Without alignments)
6657.806 Million cell updates/sec

Title: US-09-509-779-1

Perfect score: 1140

Sequence: 1 GTCCTGCGCGCGCGCCATG.....ACTAATTCATTAATTAAG 1140

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 1797656 segs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :
1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_scs:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
------------	-------	-------------	--------	-------	-------------

1	1140	100.0	1140	10	AF092877	AF092877 Mus muscu
2	1064.6	93.4	1125	10	BC011127	BC011127 Mus muscu
3	364.4	32.0	816	9	BC008627	BC008627 Homo sapi
4	362.8	31.8	836	9	AF164679	AF164679 Homo sapi
5	355.4	31.2	822	9	BC005966	BC005966 Homo sapi
6	348.4	30.6	163521	2	AC108040	AC108040 Homo sapi
7	348.4	30.6	163757	2	AC092418	AC092418 Homo sapi
8	334	29.3	754	9	AF092878	AF092878 Homo sapi
9	306.6	26.9	53155	2	AC104983	AC104983 Homo sapi
10	305.2	26.8	342	6	AF142060	AF142060 Homo sapi
11	303.6	26.6	342	6	AX282592	AX282592 Sequence
12	256.4	22.5	170087	2	AC098496	AC098496 Rattus no
13	236.6	20.8	169675	9	AC060766	AC060766 Homo sapi
14	183.4	16.1	160012	2	AC106176	AC106176 Rattus no
15	164.2	14.4	439	9	HSB333208	HSB333208 Homo sapi
16	164.2	14.4	179877	2	AC092978	AC092978 Homo sapi
17	164.2	14.3	674	2	AC067952	AC067952 Homo sapi
18	162.6	14.3	191221	2	AF312226	AF312226 Homo sapi
19	158.4	13.9	125041	2	AC068693	AC068693 Homo sapi
20	158	13.9	191221	2	AC067952	AC067952 Homo sapi
21	96	8.4	125041	2	AC068693	AC068693 Homo sapi
22	90.6	7.9	327	9	AF142059	AF142059 Homo sapi
23	90.6	7.9	497	9	HWY060A05	HWY060A05 Homo sapi
24	90.6	7.9	508	9	AF140598	AF140598 Homo sapi
25	90.6	7.9	544	9	BC001466	BC001466 Homo sapi
26	90.6	7.9	554	9	BC017370	BC017370 Homo sapi
27	90.6	7.9	5347	6	AX281690	AX281690 Sequence
28	87.8	7.7	88787	2	AC020266	AC020266 Drosophila
29	87.8	7.7	170869	3	AC011696	AC011696 Drosophila
30	87.8	7.7	171831	3	AC007473	AC007473 Drosophila
31	87.8	7.7	278196	3	AE003825	AE003825 Drosophila
32	84	7.4	553	3	AY070810	AY070810 Drosophila
33	84	7.4	750	3	AF218290	AF218290 Drosophila
34	84	7.4	3185	3	AY061302	AY061302 Drosophila
35	84	7.4	95294	2	AC019742	AC019742 Drosophila
36	84	7.4	167962	3	AC010010	AC010010 Drosophila
37	84	7.4	310120	3	AE003468	AE003468 Drosophila
38	83.8	7.4	3484	9	HSTEST	X73608 H.sapiens m
39	83.4	7.3	3208	6	AX212267	AX212267 Sequence
40	79.4	7.0	504	10	AF140599	AF140599 Mus muscu
41	76.6	6.7	389	6	AX192664	AX192664 Sequence
42	75.4	6.6	388	6	AX192698	AX192698 Sequence
43	71.2	6.2	361	3	AB077287	AB077287 Caenorhab
44	70.6	6.2	324	8	AF179228	AF179228 Schizosac
45	70.6	6.2	35277	3	CBRG19102	AC084521 Caenorhab

ALIGNMENTS

RESULT 1
LOCUS AF092877 1140 bp mRNA linear ROD 24-JUL-2001
DEFINITION Mus musculus zinc RING finger protein SAG mRNA, complete cds.
ACCESSION AF092877
VERSION AF092877.1 GI:4588031
KEYWORDS
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 1140)
Swaroop,M., Bian,J., Aviram,M., Duan,H., Blisgater,C.L., Loo,J.A.
and Sun,Y.
Expression, purification, and biochemical characterization of SAG,
a RING finger redox-sensitive protein
Free Radical Biol. Med. 27, 193-202 (1999)
TITLE
JOURNAL
AUTHORS
REFERENCE
2 (bases 1 to 1140)
Sun,Y.
Alterations of SAG mRNA in human cancer cell lines: requirement for
the RING finger domain for apoptosis protection
Carcinogenesis. 20 (10), 1899-1903 (1999)
JOURNAL
AUTHORS
REFERENCE
MEDLINE
99435944

PUBMED 10506102
 REFERENCE 3 (bases 1 to 1140)
 AUTHORS Duan,H., Wang,Y., Aviram,M., Swaroop,M., Loo,J.A., Bian,J., Tian,Y., Mueller,T., Bisgaler,C.U. and Sun,Y.
 TITLE SAA, a novel zinc RING finger protein that protects cells from apoptosis induced by redox agents
 JOURNAL Molecular and cellular biology. 19 (4), 3145-3155 (1999)
 MEDLINE 99182502
 PUBMED 10082581
 REFERENCE 4 (bases 1 to 1140)
 AUTHORS Swaroop,M., Wang,Y., Miller,P., Duan,H., Jatkoe,T., Madore,S.J. and Sun,Y.
 TITLE Yeast homolog of human SAG/ROC2/Rbx2/Hrt2 is essential for cell growth, but not for germination: chip profiling implicates its role in cell cycle regulation
 JOURNAL Oncogene. 19 (24), 2855-2866 (2000)
 MEDLINE 20309864
 PUBMED 10851089
 REFERENCE 5 (bases 1 to 1140)
 AUTHORS Kung,H.F., Zhang,H. and Sun,Y.
 TITLE Promotion of S-phase entry and cell growth under serum starvation by SAG/ROC2/Rbx2/Hrt2, an E3 ubiquitin ligase component: association with inhibition of p27 accumulation
 JOURNAL Molecular carcinogenesis. 30 (1), 37-46 (2001)
 MEDLINE 21152847
 PUBMED 1125262
 REFERENCE 6 (bases 1 to 1140)
 AUTHORS Sun,Y.
 TITLE Direct Submission
 JOURNAL Submitted (16-SEP-1998) Department of Molecular Biology, Parke-Davis, 2800 Plymouth Rd, Ann Arbor, MI 48105, USA
 FEATURES
 source 1. 1140
 /organism="Mus musculus"
 /strain="C57BL/6XCBa"
 /db_xref="taxon:10090"
 /sex="female"
 /tissue_type="lung"
 /dev_stage="6-8 weeks"
 17..358
 /note="redox sensitive, metal binding; expression protects cells from apoptosis induced by redox compounds"
 /codon_start=1
 /product="zinc RING finger protein SAG"
 /protein_id="AAD25961.1"
 /db_xref="GI:4588032"
 /translation="MADYEDDEEPCVSSHSIGSKSGGKRMFSLKKNVAVAMKMSNDVECDICARVQVMDACLRCQAKENKQEDCVVWVECHNSFHNCKMSLWVKNRNCPLC OODWVYORIGK"
 BASE COUNT 302 a 224 c 287 g 327 t
 ORIGIN
 Query Match 100.0%; Score 1140; DB 10; Length 1140;
 Best Local Similarity 100.0%; Pred. No. 1,4e-278;
 Matches 1140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 241 GGTCTGGGAGAGTGTACCAATTCCTTCCACACTGCTGATGCTCCCTGGTGAACA 300
 Db 241 GGTCTGGGAGAGTGTACCAATTCCTTCCACACTGCTGATGCTCCCTGGTGAACA 300
 QY 301 GAACAATGCTGCCCTCTGTGCCACAGAGAGTGGGTAGTCCAAAGATGGCAATGAGA 360
 Db 301 GAACAATGCTGCCCTCTGTGCCACAGAGAGTGGGTAGTCCAAAGATGGCAATGAGA 360
 QY 361 GGTGGCCAGAGCGCTCCCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 420
 Db 361 GGTGGCCAGAGCGCTCCCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 420
 QY 421 GATTCATCTTGAAGAGAGAGAGATGCTGTGCGCTTGGAGACTCACCAGAGCTTGCTT 480
 Db 421 GATTCATCTTGAAGAGAGAGAGATGCTGTGCGCTTGGAGACTCACCAGAGCTTGCTT 480
 QY 481 TATTAATTTGCTGTTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 540
 Db 481 TATTAATTTGCTGTTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 540
 QY 541 GGCCTTCTACCTCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 600
 Db 541 GGCCTTCTACCTCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 600
 QY 601 GAAATGATCTTGTGTTTATCTGTACCCAGACTGGAGACATTGTTCACAGAACATT 660
 Db 601 GAAATGATCTTGTGTTTATCTGTACCCAGACTGGAGACATTGTTCACAGAACATT 660
 QY 661 GTTGTGTTTATGCTTGAGGGTTAAAAATAGATTAACGATGTTTACGTAACAAATTA 720
 Db 661 GTTGTGTTTATGCTTGAGGGTTAAAAATAGATTAACGATGTTTACGTAACAAATTA 720
 QY 721 ATGATTAAGAAAGCGACTCTCTCTTATCTTGTGGAGAGAGAGAGAGAGAGAG 780
 Db 721 ATGATTAAGAAAGCGACTCTCTCTTATCTTGTGGAGAGAGAGAGAGAGAGAG 780
 QY 781 CACCCCTGCTCTTCTATTTGCTGTGAATAGAGATTTTAACTGCAGTCAGTAAGAGCG 840
 Db 781 CACCCCTGCTCTTCTATTTGCTGTGAATAGAGATTTTAACTGCAGTCAGTAAGAGCG 840
 QY 841 TAACGTGTGGGTAACCTATATGCGCTAACCTGCGGGTAAACGGCTTGTCTCTAC 900
 Db 841 TAACGTGTGGGTAACCTATATGCGCTAACCTGCGGGTAAACGGCTTGTCTCTAC 900
 QY 901 TTCTCATCTTTGACTTGGCCAGAGAGCTGATGTTCAACCACTTAGTTCTAAGAAC 960
 Db 901 TTCTCATCTTTGACTTGGCCAGAGAGCTGATGTTCAACCACTTAGTTCTAAGAAC 960
 QY 961 TGTTCCTGTTTTGCGCAAGGTTGTATGTTTAACTCAAAAATATTAAGTAGAA 1020
 Db 961 TGTTCCTGTTTTGCGCAAGGTTGTATGTTTAACTCAAAAATATTAAGTAGAA 1020
 QY 1021 ATGCTTACTAGTATTAACACAGTTCATATGCAAGTTTATTAATAATATGCTT 1080
 Db 1021 ATGCTTACTAGTATTAACACAGTTCATATGCAAGTTTATTAATAATATGCTT 1080
 QY 1081 TGAGTTATTAAGTTTGAATATTAATCTTAAATCATTAATCAATTAATG 1140
 Db 1081 TGAGTTATTAAGTTTGAATATTAATCTTAAATCATTAATCAATTAATG 1140
 RESULT 2
 LOCUS BC011127 1125 bp mRNA linear ROD 30-JUL-2001
 DEFINITION Mus musculus, similar to ring finger protein 7, clone MGC:19076
 ACCESSION BC011127
 VERSION BC011127.1 GI:15029807
 KEYWORDS MGC.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
AUTHORS
TITLE
JOURNAL

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1125)

Strausberg, R.
Direct Submission
Submitted (25-JUL-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK
COMMENT
NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINI)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: villalob@bcm.tmc.edu
Villalob, D.K., Luna, R.A., Hale, S.M., Huliy, S., Lu, X., Garcia,
A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W.,
Muzny, D.M., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LINI at: <http://image.llnl.gov>
Series: IRAC Plate: 24 Row: k Column: 5
This clone was selected for full length sequencing because it
passed the following selection criteria: GenomScan gene
prediction, similarity but not identity to protein.
Location/Qualifiers

FEATURES
source

1..1125
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="MGC:19076 IMAGE:4194107"
/tissue_type="Liver, normal, 5 month old male mouse."
/clone_lib="NCI-CGAP-L19"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
12..353
/codon_start=1
/product="Similar to ring finger protein 7"
/protein_id="AAH1127.1"
/db_xref="GI:15029808"
/translation="MADVEDGEPEPCVLSHSGSAGSKSGDKMFSLKKNVAVMMSMD
VECDTCALCRVQVMDACLRCQAKNKEDECVVWGECHSHFHNCSMLWVQNNRPLC
QQDWVYORIRK"
CDS

BASE COUNT 304 a 219 c 286 g 316 t
ORIGIN

Query Match 93.4% Score 1064.6 DB 10 Length 1125;
Best Local Similarity 99.18; Pred. No. 1.9e-259;
Matches 1113; Conservative 0; Mismatches 4; Indels 6; Gaps 4;
QY 6 GCGCGCCGCCATGCGAGCTGGAGAGCGGAGAAACCTGCTCTTCTCGCACT 65
DB 1 GCGCGCCGCCATGCGAGCTGGAGAGCGGAGAAACCTGCTCTTCTCGCACT 60
QY 66 CCGGAGCGAGGCTCCAGTCCGAGGCGGAGCAAGATGTTCTCTCAAGATGGAACG 125
DB 61 CCGGAGCGAGGCTCCAGTCCGAGGCGGAGCAAGATGTTCTCTCAAGATGGAACG 120
QY 126 CCGTAGCATGTGAGCTGGAGCTTGAGTGCATGCTGCTGCACTGCAAGGTTCCAGG 185
DB 121 CCGTAGCATGTGAGCTGGAGCTTGAGTGCATGCTGCTGCACTGCAAGGTTCCAGG 180
QY 186 TGATGATGCTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 245
DB 181 TGATGATGCTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 240
QY 246 GGGGAGAGTGTAACTCTTCCACAACCTGCTCATGCTGCTGCTGCTGCTGCTGCT 305
DB 241 GGGGAGAGTGTAACTCTTCCACAACCTGCTCATGCTGCTGCTGCTGCTGCTGCT 300

QY 306 ATGCGTCCCTCCTGTCGACAGAGACTGTGAGTCCAAAGATGCGAAATGAGAGGTGG 365
DB 301 ATGCGTCCCTCCTGTCGACAGAGACTGTGAGTCCAAAGATGCGAAATGAGAGGTGG 360
QY 366 CCCAGGCGCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 425
DB 361 CCCAGGCGCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
QY 426 ATCCTTAGAGAGAGAGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 485
DB 421 ATCCTTAGAGAGAGAGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
QY 486 ATTTGCTGTTATTTTGGGAAATTCCTACATTAAGTAAATTTTAAATTTGGCCT 545
DB 481 ATTTGCTGTTATTTTGGGAAATTCCTACATTAAGTAAATTTTAAATTTGGCCT 540
QY 546 TTCCCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 605
DB 541 TTCCCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
QY 606 TGATCTTTGTTTATCTGATCCACGACT--GGAACATGTTGTTACAGAAACATTTGT 662
DB 601 TGATCTTTGTTTATCTGATCCACGACTGTGGGAACATTTGTTACAGAAACATTTGT 660
QY 663 TTGTTTATGCTGAGGGTTAAATAATAGATTAAGCAATGTTACAGTAAATTAAT 722
DB 661 TTGTTTATGCTGAGGGTTAAATAATAGATTAAGCAATGTTACAGTAAATTAAT 720
QY 723 GCATTGAAAGCCGACCTCTCTTAATCTTTTGTGTTGGAGAGAGGCAAGCGAGCCA 782
DB 721 GCATTGAAAGCCGACCTCTCTTAATCTTTTGTGTTGGAGAGAGGCAAGCGAGCCA 780
QY 783 CCCGCTGCTCTTCAATTTGCTGCTGATGAGATTTTAACTGCACTGAGTGAAGGCGTA 842
DB 781 CCCGCTGCTCTTCAATTTGCTGCTGATGAGATTTTAACTGCACTGAGTGAAGGCGTA 840
QY 843 ACTGTGGGTAACTGTAAATGTGAGCTGAACGTGCGGTAACGCGCTTGTCTCTGACTT 902
DB 841 ACTGTGGGTAACTGTAAATGTGAGCTGAACGTGCGGTAACGCGCTTGTCTCTGACTT 900
QY 903 CTCATCTTTGACTTGGCCAGGAAGCTGATGTTTCAACCATTTAGTTCTTAAAGAACTG 962
DB 901 CTCATCTTTGACTTGGCCAGGAAGCTGATGTTTCAACCATTTAGTTCTTAAAGAACTG 960
QY 963 TTTTCTGTTTGGCGGAAGTGTATGTATGTTTATGCTCAAAATATTTAGTAAAT 1022
DB 961 TTTTCTGTTTGGCGGAAGTGTATGTATGTTTATGCTCAAAATATTTAGTAAAT 1018
QY 1023 GCGTTACTAGTAACTGCAAGTTCATTAATGCAATGTTTAAATAATTTGCTTTG 1082
DB 1019 GCGTTACTAGTAACTGCAAGTTCATTAATGCAATGTTTAAATAATTTGCTTTG 1077
QY 1083 AGTATTAAAGTTGATATATATCTTAAATCATTAACCTAA 1125
DB 1078 AGTATTAAAGTTGATATATATCTTAAATCATTAACCTAA 1120

RESULT 3
BC008627
LOCUS 816 bp mRNA linear PRI 12-JUL-2001
DEFINITION Homo sapiens, ring finger protein 7, clone MGC:17274 IMAGE:4177613,
mRNA, complete cds.
ACCESSION BC008627
VERSION BC008627.1 GI:14250388
KEYWORDS MGC.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 816)
AUTHORS Strausberg, R.
TITLE Direct Submission

REFERENCE
AUTHORS
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 163757)
Kaul, R.K., Olson, M.V., Raymond, C., Clendenning, J., Ivey, R.G. and
Haugen, E.D.
Direct Submission
TITLE
JOURNAL
REFERENCE
AUTHORS
Kaul, R.K., Olson, M.V., Raymond, C., Clendenning, J., Ivey, R.G. and
Haugen, E.D.
Direct Submission
TITLE
JOURNAL
COMMENT
Submitted (04-JUL-2001) Genome Center, University of Washington,
Box 352145, Seattle, WA 98195, USA
On Jul 4, 2001 this sequence version replaced gi:8705848.

Center: University of Washington Genome Center
 Center Code: UWGC
 Web site: <http://www.genome.washington.edu>
 Contact: uwgchits@u.washington.edu
 Drafting Center: BCM
 ----- Project Information -----
 Center project name: chr-3
 Center clone name: RP11-229A12 (bc0315)

```

----- Summary Statistics -----
Sequencing vector: Plasmid: 50% of reads
Sequencing vector: unknown: 50% of reads
Chemistry: Dye-terminator ET: 47% of reads
Chemistry: Dye-terminator Big Dye: 52% of reads
Assembly program: Pirap: version 0.990319
Consensus quality: 162824 bases at least Q40
Consensus quality: 163381 bases at least Q30
Consensus quality: 163468 bases at least Q20
Insert size: 163557, sum-of-contigs
Quality coverage: 7.6x in Q20 bases; sum-of-con

```

Quality coverage: 7.6x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1	15071:	contig of 15071 bp in length
* 15072	15171:	gap of unknown length
* 15172	57886:	contig of 42715 bp in length
* 57887	57986:	gap of unknown length
* 57987	163572:	contig of 105771 bp in length.

FEATURES

source

[illegible][illegible]


```

* 11894 12661: contig of 768 bp in length
* 12662 12761: gap of 100 bp
* 12762 13525: contig of 764 bp in length
* 13526 13625: gap of 100 bp
* 13626 14347: contig of 722 bp in length
* 14348 14447: gap of 100 bp
* 14448 15172: contig of 725 bp in length
* 15173 15272: gap of 100 bp
* 15273 16007: contig of 735 bp in length
* 16008 16107: gap of 100 bp
* 16108 16855: contig of 748 bp in length
* 16856 16955: gap of 100 bp
* 16956 17711: contig of 756 bp in length
* 17712 17811: gap of 100 bp
* 17812 18568: contig of 757 bp in length
* 18569 18668: gap of 100 bp
* 18669 19421: contig of 753 bp in length
* 19422 19521: gap of 100 bp
* 19522 20229: contig of 708 bp in length
* 20230 20329: gap of 100 bp
* 20330 21061: contig of 732 bp in length
* 21062 21161: gap of 100 bp
* 21162 21909: contig of 748 bp in length
* 21910 22009: gap of 100 bp
* 22010 22773: contig of 764 bp in length
* 22774 22873: gap of 100 bp
* 22874 23623: contig of 750 bp in length
* 23624 23723: gap of 100 bp
* 23724 24489: contig of 766 bp in length
* 24490 24589: gap of 100 bp
* 24590 25354: contig of 765 bp in length
* 25355 25454: gap of 100 bp
* 25455 26174: contig of 720 bp in length
* 26175 26274: gap of 100 bp
* 26275 27043: contig of 769 bp in length
* 27044 27143: gap of 100 bp
* 27144 27900: contig of 757 bp in length
* 27901 28000: gap of 100 bp
* 28001 28753: contig of 753 bp in length
* 28754 28853: gap of 100 bp
* 28854 29572: contig of 719 bp in length
* 29573 29672: gap of 100 bp
* 29673 30419: contig of 747 bp in length
* 30420 30519: gap of 100 bp
* 30520 31262: contig of 743 bp in length
* 31263 31362: gap of 100 bp
* 31363 32052: contig of 690 bp in length
* 32053 32152: gap of 100 bp
* 32153 32909: contig of 757 bp in length
* 32910 33009: gap of 100 bp
* 33010 33736: contig of 727 bp in length
* 33737 33836: gap of 100 bp
* 33837 34598: contig of 762 bp in length
* 34599 34698: gap of 100 bp
* 34699 35456: contig of 758 bp in length
* 35457 35556: gap of 100 bp
* 35557 36274: contig of 718 bp in length
* 36275 36374: gap of 100 bp
* 36375 37111: contig of 737 bp in length
* 37112 37211: gap of 100 bp
* 37212 37954: contig of 743 bp in length
* 37955 38054: gap of 100 bp
* 38055 38783: contig of 729 bp in length
* 38784 38883: gap of 100 bp
* 38884 39686: contig of 803 bp in length
* 39687 39786: gap of 100 bp
* 39787 40535: contig of 749 bp in length
* 40536 40635: gap of 100 bp
* 40636 41387: contig of 752 bp in length
* 41388 41487: gap of 100 bp
* 41488 42254: contig of 767 bp in length
* 42255 42354: gap of 100 bp
* 42355 43046: contig of 692 bp in length

```

```

* 43047 43146: gap of 100 bp
* 43147 43885: contig of 739 bp in length
* 43886 43985: gap of 100 bp
* 43986 44741: contig of 756 bp in length
* 44742 44841: gap of 100 bp
* 44842 45604: contig of 763 bp in length
* 45605 45704: gap of 100 bp
* 45705 46429: contig of 725 bp in length
* 46430 46529: gap of 100 bp
* 46530 47244: contig of 715 bp in length
* 47245 47344: gap of 100 bp
* 47345 48088: contig of 744 bp in length
* 48089 48188: gap of 100 bp
* 48189 48948: contig of 760 bp in length
* 48949 49048: gap of 100 bp
* 49049 49803: contig of 755 bp in length
* 49804 49903: gap of 100 bp
* 49904 50648: contig of 745 bp in length
* 50649 50748: gap of 100 bp
* 50749 51506: contig of 758 bp in length
* 51507 51606: gap of 100 bp
* 51607 52345: contig of 739 bp in length
* 52346 52445: gap of 100 bp
* 52446 53185: contig of 740 bp in length
* 53186 53285: gap of 100 bp
* 53286 54059: contig of 774 bp in length
* 54060 54159: gap of 100 bp
* 54160 54901: contig of 742 bp in length
* 54902 55001: gap of 100 bp
* 55002 55740: contig of 739 bp in length
* 55741 55840: gap of 100 bp
* 55841 56577: contig of 737 bp in length
* 56578 56677: gap of 100 bp
* 56678 57433: contig of 756 bp in length
* 57434 57533: gap of 100 bp
* 57534 58290: contig of 757 bp in length

```

Query Match 26.9% Score 306.6; DR 2; Length 59155;
 Best Local Similarity 77.5%; Pred. No. 1; le-66;
 Matches 424; Conservative 0; Mismatches 89; Indels 34; Gaps 3;

```

QY 62 CACGCCGGAGCGAGCGCTCCAACTGCGAGGCGACAAAGTGTCTCTCAAGAGTGG 121
Db 7636 CACTCCAGGAGCGAGCGCTCCAACTGCGAGGCGACAAAGTGTCTCAAGAGTGG 7695
QY 122 AACGCGTAGCCATGTGAGCTGCGAGCTGAGTGCATACCTGTGCGATCTGAGGTC 181
Db 7696 AACGTGCGCCCATGTGAGCTGCGAGCTGAGTGCATACCTGTGCGATCTGAGGTC 7755
QY 182 CAGGTGATGATGCTGCTGCTGCGATGTCAGTCAAGCTGAAAACAAGAGGACTGTGTG 241
Db 7756 CAGGTGATGATGCTGCTGCTGCGATGTCAGTCAAGCTGAAAACAAGAGGACTGTGTG 7815
QY 242 GTCTGGGAGAGTGTAACTTCCTCCACACTGCTGCATGCTCCCTGTGGGTAAACAG 301
Db 7816 GTCTGGGAGAGTGTAACTTCCTCCACACTGCTGCATGCTCCCTGTGGGTAAACAG 7875
QY 302 AACATTCGCTGCCCTCTGTGCCAGAGAGCTGGGAGTGCAAAAGATCGCAATGAGAG 361
Db 7876 AACATTCGCTGCCCTCTGTGCCAGAGAGCTGGGAGTGCAAAAGATCGCAATGAGAG 7935
QY 362 GTGGGCCAGGCGCTCTGCTGTGTGTGCT-----GACC 394
Db 7936 CGGTAGGAAGGCTTCTTAGTGAGTGTTCAGAGCGCTGTGATCTGTATCAAGTGC 7995
QY 395 CTGGCAAAAGACTAAACACTGCAAGGGATTCCTCTTAGAGAG-----AGAGGATGCTG 450
Db 7996 CCTTCAAAAGCTAAGAACACTCAGGGGATTTACTTCTTCAAAATGAGAGTGCATCTGT 8055
QY 451 GCGCCTTGAAGCTAACCAAGGCTTGTATTAATTGTCTGTTTATTTAGTTTGGGAAT 510
Db 8056 GGTCTTGGGACACTCAAAAGGCTTGTATTAATTGTCTGTTTATTTAGGAAT 8115

```

QY 511 TCCTACAAATT---AGATATATTTGTTAAATGCGCTTCCACCTCTGGTGTGTGT 567
 |||||
 Db 8116 TCCTACAAATTAAAGATATATTTAAAGATGATGCTTTCACCTCTGTGATGTGT 8175
 QY 568 GTGATAC 574
 |||
 Db 8176 CGCACAC 8182

RESULT 10
 AF142060 342 bp mRNA linear PRI 14-JUL-1999
 LOCUS Homo sapiens RING finger protein (ROC2) mRNA, complete cds.
 DEFINITION AF142060
 ACCESSION AF142060.1 GI:4809217
 VERSION
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 AUTHORS Ohta,T., Michel,J.J., Schotellus,A.J. and Xiong,Y.
 TITLE ROC1, a homolog of APC11, represents a family of cullin partners
 JOURNAL Mol. Cell 3 (4), 535-541 (1999)
 MEDLINE 99247022
 PUBMED 10230407
 REFERENCE
 AUTHORS Ohta,T., Michel,J.J. and Xiong,Y.
 TITLE Direct Submission
 JOURNAL Submitted (10-APR-1999) Lineberger Comprehensive Cancer Center,
 University of North Carolina at Chapel Hill, Mason Farm Rd. and
 Manning Dr., Chapel Hill, NC 27599-7295, USA

FEATURES

source 1..342
 location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 1..342
 /gene="ROC2"
 1..342
 /gene="ROC2"
 /note="Cullin partnering protein with associated ubiquitin
 ligase activity"
 /codon_start=1
 /product="RING finger protein"
 /protein_id="AA030147.1"
 /db_xref="GI:4809218"
 /translation="MADVEDGEETCALASHSGSGSTSGDKMFLKKNVAVAMMSMD
 VEDICAIKRVQVMDACLRCAENKQEDCVVVGECNHSFHNCMSLMWKONNRCPDLC
 OODMYYQIRIGK"
 OODMYYQIRIGK"

BASE COUNT 82 a 83 c 109 g 68 t

ORIGIN

Query Match

Best Local Similarity 26.8%; Score 305.2; DB 9; Length 342;
 Matches 319; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 17 ATGGCCGACGTGAGAGCGGCGAGGAACCTGCGCTCTTCTTCTGCACTCCGGAGCGCA 76
 |||||
 Db 1 ATGGCCGACGTGAGAGCGGCGAGGAACCTGCGCTCTTCTGCACTCCGGAGCGCA 60
 QY 77 GGGTCCCAAGTCGGGAGCGGCGAGGATGCTCTCTCAAGAGTGGAGCGGGTGCATG 136
 |||||
 Db 61 GGGTCCCAAGTCGGGAGCGGCGAGGATGCTCTCTCAAGAGTGGAGCGGGTGCATG 120
 QY 137 TGGAGCTGGAGCGTTGAGTGCATCTGCGCATCTGCAAGGGTCCAGGTGATGATGCC 196
 |||||
 Db 121 TGGAGCTGGAGCGTTGAGTGCATCTGCGCATCTGCAAGGGTCCAGGTGATGATGCC 180
 QY 197 TGCCTTGATGTCAAGCTGAAGCAAGAGAGAGCTGTGTGTGCTGGGAGAGAGTGT 256
 |||||
 Db 181 TGTCTTAGATGTCAAGCTGAAGCAAGAGAGAGAGCTGTGTGTGCTGGGAGAGAGTGT 240

QY 257 ACCATTCTTCCCAACTGCTGATGCTCCCTGTGGGTGAAGAGCAATCGCTGCCCT 316
 |||||
 Db 241 AATCATTTCTTCCCAACTGCTGATGCTCCCTGTGGGTGAAGAGCAATCGCTGCCCT 300
 QY 317 CTGTGCCAGCAGGACTGGTGTAGTCCAAAGATGGCAATGA 358
 |||||
 Db 301 CTGTGCCAGCAGGACTGGTGTAGTCCAAAGATGGCAATGA 342

RESULT 11

AX282592 342 bp DNA linear PAT 02-NOV-2001
 LOCUS AX282592
 DEFINITION Sequence 7 from Patent WO0175145.
 ACCESSION AX282592
 VERSION AX282592.1 GI:16609675
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 AUTHORS Issakant,S.D., Huang,J., Sheung,J. and Pray,T.R.
 TITLE Ubiquitin ligase assay
 JOURNAL Patent: WO 0175145-A 7 11-OCT-2001;
 Rigel Pharmaceuticals, Inc. (US)
 location/Qualifiers
 1..342
 /organism="Homo sapiens"
 /db_xref="taxon:9606"

BASE COUNT 82 a 84 c 108 g 68 t

ORIGIN

Query Match

Best Local Similarity 26.8%; Score 303.6; DB 6; Length 342;
 Matches 318; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 17 ATGGCCGACGTGAGAGCGGCGAGGAACCTGCGCTCTTCTTCTGCACTCCGGAGCGCA 76
 |||||
 Db 1 ATGGCCGACGTGAGAGCGGCGAGGAACCTGCGCTCTTCTGCACTCCGGAGCGCA 60
 QY 77 GGGTCCCAAGTCGGGAGCGGCGAGGATGCTCTCTCAAGAGTGGAGCGGGTGCATG 136
 |||||
 Db 61 GGGTCCCAAGTCGGGAGCGGCGAGGATGCTCTCTCAAGAGTGGAGCGGGTGCATG 120
 QY 137 TGGAGCTGGAGCGTTGAGTGCATCTGCGCATCTGCAAGGGTCCAGGTGATGATGCC 196
 |||||
 Db 121 TGGAGCTGGAGCGTTGAGTGCATCTGCGCATCTGCAAGGGTCCAGGTGATGATGCC 180
 QY 197 TGCCTTGATGTCAAGCTGAAGCAAGAGAGAGCTGTGTGTGCTGGGAGAGTGT 256
 |||||
 Db 181 TGTCTTAGATGTCAAGCTGAAGCAAGAGAGAGCTGTGTGTGCTGGGAGAGTGT 240
 QY 257 AACCATCTTCCCAACTGCTGATGCTCCCTGTGGGTGAAGAGCAATCGCTGCCCT 316
 |||||
 Db 241 AATCATTTCTTCCCAACTGCTGATGCTCCCTGTGGGTGAAGAGCAATCGCTGCCCT 300
 QY 317 CTGTGCCAGCAGGACTGGTGTAGTCCAAAGATGGCAATGA 358
 |||||
 Db 301 CTGTGCCAGCAGGACTGGTGTAGTCCAAAGATGGCAATGA 342

RESULT 12

AC098496 170087 bp DNA linear HTG 20-DEC-2001
 LOCUS AC098496/c
 DEFINITION Rattus norvegicus clone CH230-23c21, *** SEQUENCING IN PROGRESS
 *** 62 unordered pieces.

AC098496.3 GI:17973133
 VERSION AC098496.3
 KEYWORDS HTG; HTGS_PHASEL.
 SOURCE Norway rat.
 ORGANISM Rattus norvegicus

REFERENCE
AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

1 (bases 1 to 170087)

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-osman, F.R., Allen, C.,
Alsbrooks, S.L., Amaral, H.C., Are, J.R., Banks, T., Barbata, J.,
Benton, J., Bimaga, K., Blankenburg, K., Bonin, D., Bouck, J.,
Bowie, S., Brieau, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C.,
Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R.,
Chen, Z., Chowdhury, I., Christopoulos, C., Cleveland, C.D., Cox, C.,
Coyle, M.D., Dethorne, S.R., David, R., Davila, M.L., Davis, C.,
Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O.,
Denn, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Diaper, H.,
Dugan-Rocha, S., Durbin, K.J., Earnhart, C., Edgar, D., Edwards, C.C.,
Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J.,
Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T.,
Garza, N., Gill, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S.,
Hamilton, K., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A.,
Hernandez, J., Hernandez, O., Hodgson, A., Hogues, M., Hollway, C.,
Hollins, B., Homs, F., Howard, S., Huber, J., Huliy, S., Hume, J.,
Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jollivet, S.,
Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J.,
Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C.,
Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W.,
Lounsbeg, H., Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R.,
Ma, J., Maheshwari, M., Mapua, P., Martin, R., Marindale, A.,
Martinez, E., Massey, E., Mawhinney, E., McLeod, M.P., Meador, M.,
Mei, G., Metzger, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K.,
Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N.,
Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwankwo, S.,
Ogulu, M., Okunolu, G., Oragunye, N., Oviedo, R., Pace, A., Pu, L.,
Peety, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.,
Quiles, S., Saverly, G., Scherer, S., Scott, G., Shen, H., Shostkari, N.,
Sisson, I., Sodergren, E., Sonalke, T., Sparks, A., Stanley, H.,
Stane, H., Sutton, A., Svatek, A., Tabor, P., Tameis, A., Tameis, K.,
Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N.,
Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalob, D., Vinson, R.,
Wall, R., Wang, S., Ward-Moore, S., Warren, R., Washington, C.,
Washington, S., Williams, G., Williamson, A., Wleczko, R., Wooden, S.,
Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorilla, S., Nelson, D.,
Weinstock, G. and Gibbs, R.

Unpublished

2 (bases 1 to 170087)

Worley, K.C.

Direct Submission

Submitted (24-OCT-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Dec 20, 2001 this sequence version replaced gi:17062512.

Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

Project Information

Center project name: CHEY

Center clone name: CH230-23C21

Summary Statistics

Assembly program: Phrap; version 0.990329First call to

findPrphrap

Consensus quality: 130294 bases at least 040

Consensus quality: 139538 bases at least 030

Consensus quality: 146952 bases at least 020

Estimated insert size: 136372; sum-of-contigs estimation

Quality coverage: 0x in Q20 bases; agarose-gel estimation

Quality coverage: 2.1x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length

* (see http://www.hgsc.bcm.tmc.edu/docs/denbank_draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently

* consists of 62 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1	6781:	contig of 6781 bp in length
*	6782	gap of unknown length
*	6882	contig of 5054 bp in length
*	11935:	gap of unknown length
*	11936	contig of 4708 bp in length
*	12035:	gap of unknown length
*	12036	contig of 4708 bp in length
*	16743:	gap of unknown length
*	16744	contig of 5663 bp in length
*	22506:	gap of unknown length
*	22607	contig of 7175 bp in length
*	22607	contig of 7175 bp in length
*	29782	gap of unknown length
*	29882	contig of 5688 bp in length
*	35570	gap of unknown length
*	35669:	gap of unknown length
*	38472:	contig of 2803 bp in length
*	38473	gap of unknown length
*	38572:	gap of unknown length
*	44500:	contig of 5928 bp in length
*	44501	gap of unknown length
*	44601	contig of 5196 bp in length
*	49795:	gap of unknown length
*	49896:	gap of unknown length
*	51963:	contig of 2067 bp in length
*	52063:	gap of unknown length
*	54511:	contig of 2448 bp in length
*	54512	gap of unknown length
*	54611:	contig of 3828 bp in length
*	58439:	gap of unknown length
*	58440	contig of 4365 bp in length
*	62904:	gap of unknown length
*	62905	contig of 2360 bp in length
*	63005	gap of unknown length
*	63365:	gap of unknown length
*	65465	contig of 3337 bp in length
*	68801:	gap of unknown length
*	68802	contig of 2165 bp in length
*	68902	gap of unknown length
*	71067	contig of 3904 bp in length
*	71167	gap of unknown length
*	75070:	contig of 3428 bp in length
*	75071	gap of unknown length
*	75171	contig of 4119 bp in length
*	78599	gap of unknown length
*	78699	contig of 4119 bp in length
*	82817:	gap of unknown length
*	82818	contig of 3200 bp in length
*	86117:	gap of unknown length
*	86217:	contig of 2252 bp in length
*	86218	gap of unknown length
*	88470	contig of 1730 bp in length
*	88570	gap of unknown length
*	90299:	contig of 2099 bp in length
*	90300	gap of unknown length
*	90399:	contig of 2099 bp in length
*	92498:	gap of unknown length
*	92499	contig of 3110 bp in length
*	92599	gap of unknown length
*	95709	contig of 2641 bp in length
*	95809:	gap of unknown length
*	98449:	contig of 2641 bp in length
*	98450	gap of unknown length
*	98549:	contig of 2154 bp in length
*	98550	gap of unknown length
*	100703:	contig of 2389 bp in length
*	100704	gap of unknown length
*	100803:	contig of 3452 bp in length
*	100804	gap of unknown length
*	103393:	contig of 2704 bp in length
*	103492:	gap of unknown length
*	103493	contig of 2704 bp in length
*	106944:	gap of unknown length
*	107044:	contig of 2704 bp in length
*	107045	gap of unknown length
*	109749	contig of 2657 bp in length
*	109849	gap of unknown length
*	112505:	contig of 3570 bp in length
*	112605:	gap of unknown length
*	116175:	contig of 1647 bp in length
*	116176	gap of unknown length
*	117922:	contig of 1647 bp in length
*	117923	gap of unknown length
*	118022:	contig of 2049 bp in length
*	118023	gap of unknown length
*	120071:	contig of 2049 bp in length
*	120072	gap of unknown length
*	120171:	gap of unknown length

Query Match	22.5%;	Score 256.4;	DB 2;	Length 170087;
Best Local Similarity	64.8%;	Pred. NO. 7.3e-54;		
Matches 653; Conservative	0;	Mismatches 242;	Indels 112;	Gaps 14

D	47737	TCATTAATTAGATGATTTTATT - AAAATAAGCCTTCCACCTCTGTGGTGCTGTGATAT	47679
OY	573	ACGAATGCATGAGAGCGAGACACCAGAAAATGATCTTTGTTATCTGTACCCA ---	628
D	47678	GTAGCTGTTTTAACCCAGCAATGCAGAGATCTTTGTTGTTATCTGTACATAAAGG	47619
OY	629	CGACTGGAAACAATTGTCACACAGAACAATTGTTGTTATATGCTTGAGGGTTAAAAA	688
D	47618	CTGGGAGAGCATTTGGCTTCACAGAAAGACATTTGGTGT - TTAATGCTTGGCTATAAAAA	47560
OY	689	ATAGATAAACAAGATTGTCAGTAAACAAATAAATGCATTGAAAAAGCCGACTCTCC ---	744
D	47559	ATAAATAAATAAATGTACACTAAAAAATTAATGCATTGAGAACCAACTCTCACTT	47500
OY	745	-----TAATCCTTTTGTGTGGGAGAGAGCAAGGAGGCCACCCGTGCTTGAT	797
D	47499	TGCTCATATATACCCCTTTTGTGTGGAGCTGAGGCAACCAATGTAATCTTTCTATATA	47440
OY	798	TTG-----CTGTGAATGAGATTTTAACCTG	823
D	47439	ATGAGGGTTAACTTGTACTGAGTGAAGAGGGTAATTCAGATGTAACTGTTAAGAAACAG	47380
OY	824	CACATCAGTGAAGAGCGCTAATCTGTGGGTAA -----CTG	858
D	47379	AAGTCGGGCTGGGGATTTAGCTCACTGGGTAGAGCGCTTCCTTAGGAACGCCAGGCCCTG	47320
OY	859	TAATATGCGTAACACTGTGGGTAAACGGCTTTGTCTCGACTTCTCCATCTTTGACTTG	918
D	47319	GGTTCGTCCTCCAGCTCCGAAAAAAAAAGACCCAAAAAAAAAAAAAAAAAAGAACCAAG	47260
OY	919	GCCAGGAGC----CTGGATTTGTYCAACCACTAGTCTTAAGAAGACTTTTCTGTTTT	973
D	47259	TCAAAGAAACTACTGTTCACGTTCACACCGCTTTTGAAGGATTAAGAAGATTTCTGTTTT	47200
OY	974	TGCGAAGGTTGATATGATGTTTATGTATAGCAAAAAATATATAGTAGGAATAATGGCTTACTAGT	1033
D	47199	TG-GGAAGTTGTATTTGATGTTTATGTC -AAAAATTTATGTGGAAAAAT- GCTTAGCACT	47143
OY	1034	ATAACACTGAAGTTCATTATGCAATGTTTTATA---AAATATGTGCTTTGAGTTATT	1089
D	47142	ATGGCACCGGAGTTCATTATGCAATGCTTTAATAATGAATAATATGTGCTTTGAGTTATT	47083
OY	1090	AAAGTTTATATATCTCTTAAAAATCAATTAACATATCATCATTA 1136	
D	47082	AAAGTTGATCTCTACTCTTAAATTCATTAATAATATTCAGTAATA 47036	

```

Levine, R., Lieu, C., Liu, G., Locke, K., MacDonald, P., Margulis, N.,
McCarthy, M., McKean, P., McGurk, A., McKernan, K., McPheeters, R.,
Meldrim, J., Menus, L., Mihova, T., Miranda, C., Mlenga, V., Morrow, J.,
Murphy, T., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P.,
O'Neill, D., Oliver, T. M., Oliver, J., Peterson, K., Pierre, N., Rothman, D.,
Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Roy, A.,
Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
Tesfaye, S., Theodore, J., Tirrell, A., Travers, M., Triggillo, J.,
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J.,
Young, G., Zainoun, J., Zimmer, A. and Zody, M.

TITLE
JOURNAL
Submitted (20-APR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 169675)
REFERENCE
AUTHORS
Birken, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Boguski, L., Bouckgeater, B.,
Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B.,
Chepel, T., Colangelo, M., Collins, S., Collymore, A., Cook, A.,
Cooke, P., Deavellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S.,
Ferrelira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,
Gilde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,
Jones, C., Kamat, A., Karatas, A., Kells, C., Labrecque, K.,
Lamazares, R., Landers, T., Lehotzky, J., Levine, R., Liu, G.,
Maclean, C., MacDonald, P., Major, J., Margulis, N., Matthews, C.,
McCarthy, M., McKean, P., McKernan, K., McPheeters, R., Meldrim, J.,
Menus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C.,
Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D.,
Oliver, J., Peterson, K., Phunhang, P., Pierre, N., Pollara, V.,
Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
Roman, J., Roselli, M., Roy, A., Santos, R., Schauer, S., Schupack, R.,
Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
Topham, K., Travers, M., Travis, N., Triggillo, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (01-DEC-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Nov 15, 2001 this sequence version replaced gi:15422023.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
Center: Whitehead Institute/MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information
Center Project name: L6276
Center Clone name: 686_D_22

-----
FEATURES
source
Location/Qualifiers
1..169675
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="17"
/map="17"
/clone="RP11-686D22"
/clone_lib="RPCT-11 Human Male BAC"
239..259
/rpt_family="AT-rich"
263..509
/rpt_family="AluYo"
complement(618..1197)
/rpt_family="L1MC4"
1198..1218
/rpt_family="(T5)n"
1758..1933
/rpt_family="AluYo"
1934..2116
/rpt_family="L1"
/rpt_family="L1M4"

repeat_region 2123..2237
/rpt_family="AluYo"
repeat_region 2238..2262
/rpt_family="(TAAA)n"
repeat_region 2263..2367
/rpt_family="AluYo"
repeat_region complement(2796..3097)
/rpt_family="AluYo"
repeat_region 3106..3161
/rpt_family="CT-rich"
repeat_region 3162..3338
/rpt_family="(TTC)n"
repeat_region complement(3354..3672)
/rpt_family="AluSq"
repeat_region 4976..5489
/rpt_family="MER51A"
repeat_region 5490..5664
/rpt_family="MER51A-int"
repeat_region 5656..6452
/rpt_family="MER51A-int"
repeat_region 6452..6747
/rpt_family="MER57-int"
repeat_region 6746..7310
/rpt_family="MER57-int"
repeat_region 7314..7401
/rpt_family="LFR54B"
repeat_region 7402..7491
/rpt_family="MER57-int"
repeat_region 9428..9535
/rpt_family="LIME"
repeat_region 9925..10030
/rpt_family="LIME"
repeat_region 10325..10387
/rpt_family="L1M4C"
repeat_region 10819..11180
/rpt_family="L1M4C"
repeat_region 11178..12391
/rpt_family="L1MA8"
repeat_region complement(12392..12694)
/rpt_family="AluYo"
repeat_region 12695..12868
/rpt_family="L1MA8"
repeat_region 13176..13320
/rpt_family="(CTTA)n"
repeat_region complement(13647..13998)
/rpt_family="MLT1A1"
repeat_region 15861..16165
/rpt_family="AluSc"
repeat_region complement(16442..16773)
/rpt_family="AluX"
repeat_region complement(17624..17785)
/rpt_family="MLT1K"
repeat_region complement(17816..18567)
/rpt_family="LIME"
repeat_region 18588..18815
/rpt_family="AluX"
repeat_region 18822..18858
/rpt_family="(CAA)n"
repeat_region complement(19014..19416)
/rpt_family="T199er2a"
repeat_region complement(19486..20166)
/rpt_family="LIME"
repeat_region 20167..20251
/rpt_family="L1PA6"
repeat_region complement(20252..20764)
/rpt_family="LIME"
repeat_region complement(20773..20943)
/rpt_family="AluYo"
repeat_region complement(21195..23103)
/rpt_family="L1PB1"
repeat_region complement(23105..23925)
/rpt_family="L1"
repeat_region complement(23926..24234)

```

Query Match	20.8%;	Score 236.6;	DB 9;	Length 169675;
Best Local Similarity	68.5%;	Pred. NO. 7.8e-49;		
Matches 416;	Conservative 0;	Mismatches 154;	Indels 37;	Gaps 5

Db 92891 TCCTACAAATTAAGAGATTAATTACTACAAAGAGCTTTCTTCTACCTGAGGTGTGCT 92892

Ob 568 GTGATAC 574
11
Db 92891 CACACAC 92885

RESULT 14
AC106176/c
LOCUS
DEFINITION
AC106176 160012 bp DNA linear HTG 12-JAN-2002
Rattus norvegicus clone CH230-20P16, *** SEQUENCING IN PROGRESS
***, 77 unordered pieces.
AC106176
AC106176.1 GI:18138697
HTG: HTG6.PHASE1.
SOURCE
Rattus norvegicus
Norway rat.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE 1 (bases 1 to 160012)
Muzny,D.M., Adams,C., Adio-Ogunola,B., Ali-osman,F.R., Allen,C.,
Alstbrooke,S.L., Amaralunge,H.C., Ate,J.R., Banks,T., Barbarta,J.,
Benton,J., Bimuge,K., Blankenburg,K., Bonin,D., Bouck,J.,
Bowrie,S., Brieva,M., Brown,E., Brown,M., Bryant,A.P., Buhay,C.,
Burich,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carion,T.F.,
Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
Chen,Z., Chowdhury,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
Covley,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,
Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagay,N., Ford,J.,
Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
Hamilton,K., Harris,C., Harris,K., Hart,M., Havlik,P., Hawes,A.,
Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C.,
Hollins,B., Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,
Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
Joudan,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
Lewis,L., Li,J., Li,Z., Licharge,O., Lieu,C., Liu,J., Liu,W.,
Louliseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
Ma,J., Maheshwari,M., Mapa,P., Martin,R., Martindale,A.,
Martinez,E., Massey,E., Mawhney,E., McLeod,M.P., Meador,M.,
Mel-G., Metker,M., Miner,G., Miner,E., Mitchell,T., Mohabbat,K.,
Morgen,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N.,
Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Norkenhu,S.,
Ogunu,M., Okunonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,
Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,
Qualls,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M.,
Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shoochert,I.,
Slisson,I., Sodeguren,E., Sonike,T., Sparks,A., Stanley,H.,
Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K.,
Thang,H., Tansey,J., Taylor,C., Taylor,R., Telifod,B., Thomas,N.,
Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vilson,R.,
Wali,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,
Wellington,S., Williams,G., Williams,A., Wlezyk,R., Wooden,S.,
Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Unpublished
2 (bases 1 to 160012)
Morley,K.C.
Direct Submission
Submitted (12-JAN-2002) Human Genome Sequencing Center, Department
Of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 25, 2002, 19:07:32 ; Search time 102.69 Seconds
(without alignments)
1803.561 Million cell updates/sec

Title: US-09-509-779-3

Perfect score: 754
Sequence: 1 ATGGCCGACGCGGAAGACG.....CAATTAAGTCACTTTAA 754

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 38353 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

Issued Patents_NA: *
1: /cgn2_6/ptodata/2/lna/5A.COMB.seq: *
2: /cgn2_6/ptodata/2/lna/5B.COMB.seq: *
3: /cgn2_6/ptodata/2/lna/6A.COMB.seq: *
4: /cgn2_6/ptodata/2/lna/6B.COMB.seq: *
5: /cgn2_6/ptodata/2/lna/PTOTUS.COMB.seq: *
6: /cgn2_6/ptodata/2/lna/Backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	36	4.8	1342	3	US-08-832-399-1
C 2	36	4.8	1342	4	US-09-372-498-1
C 3	36	4.8	1575	3	US-08-858-876A-1
C 4	36	4.8	1575	4	US-09-472-880-1
5	35.6	4.7	176373	3	US-09-128-155-17
6	35.2	4.7	716	4	US-08-998-416-746
7	34.2	4.5	433	4	US-09-040-984-58
8	34.2	4.5	433	4	US-09-123-912-58
C 9	33.2	4.4	302	4	US-09-656-286-10
C 10	33.2	4.4	302	4	US-09-656-286-12
C 11	33.2	4.4	423	1	US-08-470-179-55
C 12	32.2	4.3	1779	4	US-09-371-696-1
C 13	31.6	4.2	302	4	US-09-656-286-11
C 14	31.6	4.2	302	4	US-09-656-286-13
C 15	31.4	4.2	423	1	US-08-470-179-142
C 16	31.4	4.2	3404	1	US-08-265-429A-1
17	31.4	4.2	3404	5	PCT-US95-09069-1
C 18	31.4	4.2	7218	1	US-08-232-463-14
C 19	31.2	4.1	1631	3	US-09-118-319-1
20	31	4.1	4195	3	US-08-340-011-1
21	31	4.1	4416	3	US-08-901-710-1
22	31	4.1	4416	3	US-08-795-430-1
23	31	4.1	4425	1	US-08-222-616-31
24	31	4.1	4425	5	PCT-US95-04228-31
25	31	4.1	4795	1	US-08-340-011-3
26	31	4.1	4795	3	US-08-901-710-3
27	31	4.1	9108	5	PCT-US95-04228-45

C 28	31	4.1	35081	2	US-08-752-760A-1	Sequence 1, Appl
C 29	30.8	4.1	1903	3	US-08-581-148C-15	Sequence 15, Appl
C 30	30.8	4.1	6924	1	US-08-015-973-2	Sequence 2, Appl
31	30.8	4.1	6924	2	US-08-448-164-2	Sequence 2, Appl
32	30.6	4.1	246240	2	US-08-724-394A-20	Sequence 20, Appl
33	30.6	4.1	246240	2	US-08-724-394A-21	Sequence 21, Appl
34	30.6	4.1	246240	2	US-08-724-394A-22	Sequence 22, Appl
C 35	30.4	4.0	1185	3	US-09-023-339-3	Sequence 3, Appl
C 36	30.4	4.0	1260	3	US-09-023-173-5	Sequence 5, Appl
C 37	30.4	4.0	1308	3	US-09-023-339-6	Sequence 10, Appl
C 38	30.2	4.0	791	4	US-08-858-207A-219	Sequence 219, Appl
39	30.2	4.0	4451	3	US-08-717-294-42	Sequence 42, Appl
C 40	30.2	4.0	302	3	US-09-154-083-5	Sequence 5, Appl
C 41	30	4.0	302	4	US-09-656-286-14	Sequence 14, Appl
C 42	30	4.0	302	4	US-09-656-286-15	Sequence 15, Appl
C 43	30	4.0	400	1	US-08-644-326-1	Sequence 1, Appl
C 44	30	4.0	400	1	US-09-022-238-1	Sequence 1, Appl
C 45	30	4.0	400	4		

ALIGNMENTS

RESULT 1
US-08-832-399-1/C
Sequence 1, Application US/08832399
Patent No. 6008050
GENERAL INFORMATION:
APPLICANT: Bergsma, Dirk
APPLICANT: Shabon, Usman
TITLE OF INVENTION: NOVEL HUMAN NEURENIN RECEPTOR TYPE 2
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY:
ZIP: 19406
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/832,399
FILING DATE: 02-APR-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: King, William T
REGISTRATION NUMBER: 30,954
REFERENCE/DOCKET NUMBER: GH50020
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-5515
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1342 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-832-399-1

Query Match 4.8%; Score 36; DB 3; Length 1342;
Best Local Similarity 54.5%; Pred. No. 0.11;
Matches 72; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

Oy	28	ACCTGGCGCCCTGGGCTCTCATCTCCGGAGCTCAGGCTTCACAAATGGGAGCGACCAAGATG	87
Db	645	ACCAAGCAACGTCGACACTCGCGGAGGGGGCTCCGGCTCCCGCTCCGCGCTCTGAGATTGC	586
Oy	88	TTTCCCTCCCAAGAGTGGAAACCGCTGGCGCCATGTGGAGCTGGGACGTGGAGTGTCCGATACG	147
Db	565	TGCTTCTTCCGCCCATGATGACCGGCTCATGGGACGAGGCCGAGGCCGAGCGGCCCAAGAG	526
Oy	148	TGGCGCATCTGC	159
Db	525	AGCGGCACCAAGC	514

```

RESULT 2
US-09-372-498-1/c
: Sequence 1, Application US/09372498
: Patient No. 6166182
: GENERAL INFORMATION:
: APPLICANT: Derk J. Bergsma
: APPLICANT: Usman Shabon
: TITLE OF INVENTION: NOVEL HUMAN NEUTROGENSIN RECEPTOR TYPE 2
: TITLE OF INVENTION: AND SPLICED VARIANTS THEREOF
: FILE REFERENCE: GH-50020-1
: CURRENT APPLICATION NUMBER: US/09/372,498
: CURRENT FILING DATE: 1999-08-11
: PRIOR APPLICATION NUMBER: 08/832,399
: PRIOR FILING DATE: 1997-04-02
: NUMBER OF SEQ ID NOS: 12
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 1
: LENGTH: 1342
: TYPE: DNA
: ORGANISM: HOMO SAPIENS
: US-09-372-498-1

```

	Query Match	4.8%	Score 36	DB 4	Length 1342
	Best Local Similarity	54.5%	Pred. NO. 0.11		
	Matches 72	Conservative 0	Mismatches 60	Indels 0	Gaps 0
QY	28	ACCTGCGCCCTGAGGCTCTTCACATCCCGGAGACTGAGGCTCCCAATCGGAGGCGACAGATG	87		
Db	645	ACCAAGCACCGTGCACACACACGCGGAGGCGGCGTCCGGCTCCCGCTTCGAGATTGG	586		
QY	88	TTTCTCCTCAGAGAACTGCAACGCGGTGGCCATGTGAGCTGGAGCTGGAGTGGCGATTAC	147		
Db	585	TGCTTCTCTCCCATGATGATGAGCGCATGGGAGGCGAGGCGGAGGCGGCGCCACAGAG	526		
QY	148	TGGCGCATCTGC	159		
Db	525	AGCGCCACCAAGC	514		

RESULT 3
 US-08-858-876A-1/c
 : Sequence 1, Application US/08858876A
 : Patent No. 6022856
 :
 : GENERAL INFORMATION:
 :
 : APPLICANT: Daniel CAPT
 : APPLICANT: Pascal CHALON
 : APPLICANT: Pascual FERRARA
 : APPLICANT: Vita METALIO
 :
 : TITLE OF INVENTION: Type 2 Neurotensin Receptor
 : TITLE OF INVENTION: (hnt-R2)
 : NUMBER OF SEQUENCES: 12
 :
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Jacobson, Price, Holman & Stern, PLLC
 : STREET: 400 Seventh Street
 : CITY: Washington
 : STATE: D.C.
 : COUNTRY: USA
 : ZIP: 20004
 :
 : COMPUTER READABLE FORM:

```

? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/858,876A
? FILING DATE: 19-SEP-1997
? CLASSIFICATION: 536
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: PCT/FR 9723204
? FILING DATE: 17-MAR-1997
? ATTORNEY/AGENT INFORMATION:
? NAME: Player, William E.
? REGISTRATION NUMBER: 31,049
? INFORMATION FOR SEQ ID NO: 1:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1575 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: Linear
? MOLECULE TYPE: cDNA
? ORIGINAL SOURCE:
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 37..1266
? US-08-858-876A-1

```

Query Match	4.8%;	Score 36;	DB 3;	Length 1575;
Best Local Similarity	54.5%;	Pred. No. 0.12;		
Matches	72;	Conservative	0;	Mismatches 60;
			Indels	0;
			Gaps	0;

QY	28	ACCTGCGCCCTGGGCTCTCATCTCGGGAGCTCAGGCTCTCAAGTCGGGAGGCGACCAAGATG	87
Db	629	ACCAAGCAACGTGCACACTCGGAGAGGGGGGCTCGCGCTCTCCCGTCTCGCGCTCTGAGATTGG	570
QY	88	TTTCTCCCTCAAGAAAGTGGAAACGGGTGGCCATGTGGAGCTGGGACGTGGAGTGGCATACG	147
Db	569	TGCTTCTTCCGCCCATGATGATACGGCCATGGGACGAGGCGCGAACCGAGGGCGGCCACAGAG	510
QY	148	TGGGCAATCTTG	159
Db	509	AGCGCCACCAACG	498

RESULT 4
US-09-472-880-1/C
Sequence 1, Application US/09472880
Patent No. 6274333
GENERAL INFORMATION:
APPLICANT: Daniel CAPUT
Pascale CHALON
Pascale FERRARA
Vita NATALIO
TITLE OF INVENTION: Type 2 Neurotensin Receptor
(hMT-R2)
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Jacobson, Price, Holman & Stern, PLLC
STREET: 400 Seventh Street
City: Washington
STATE: D. C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/472,880
FILING DATE: 28-Dec-1999

CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FR 9723204
FILING DATE: 17-MAR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Player, William E.
REGISTRATION NUMBER: 31,049
INFORMATION FOR SEQ ID NO: 1
SEQUENCE CHARACTERISTICS:
LENGTH: 1575 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: 37..1266
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-472-880-1

Query Match 4.8%; Score 36; DB 4; Length 1575;
Best Local Similarity 54.5%; Pred. No. 0.12;
Matches 72; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY 28 ACCTGGCCCTGGCTCTACCTCCAGCTCCAGCGGAGCGCAGATG 87
DB 629 ACCAGACCGTGCACATCGGAGGGGCTCCCGCTCCCGCTTCGATTGC 570
QY 88 TTCTCCCTCAAGAGTGAACGGGTGGCCATGTGAGCTGGAGCGTGGAGTGCATAG 147
DB 569 TGCTTCTGCCCATGTATGACGGCATGGGAGGCGAGCGCGCCAGGCGCCACGAG 510
QY 148 TGGCCCATCTGC 159
DB 509 AGCGCCACGAGC 498

RESULT 5
US-09-128-155-17
Sequence 17, Application US/09128155
Patent No. 6117654
GENERAL INFORMATION:
APPLICANT: Pan, Yang
TITLE OF INVENTION: NOVEL MOLECULES OF TANGO-77 RELATED PROTEIN FAMILY
TITLE OF INVENTION: AND USES THEREOF
FILE REFERENCE: 09404/052001
CURRENT APPLICATION NUMBER: US/09/128,155
CURRENT FILING DATE: 1998-08-03
EARLIER APPLICATION NUMBER: US 60/091,650
EARLIER FILING DATE: 1998-07-02
EARLIER APPLICATION NUMBER: US 60/054,646
EARLIER FILING DATE: 1997-08-04
NUMBER OF SEQ ID NOS: 18
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 17
LENGTH: 176373
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(176373)
OTHER INFORMATION: n = A,T,C or G
US-09-128-155-17

Query Match 4.7%; Score 35.6; DB 3; Length 176373;
Best Local Similarity 50.0%; Pred. No. 3.1;
Matches 89; Conservative 0; Mismatches 89; Indels 0; Gaps 0;
QY 367 CAGTTGTCAGAGCCCTGGTGAATCTTGAATCAGTCCCTACAAAGCTAGAACATCA 426

DB 164426 caggtgtgtgacctcctagtcagaggtctcttgcacctgcgcccgaggtgtctaactccta 164485
QY 427 CAGGGATGAATTTCTTCAAATAGAGCCGATGATGTGTGGCTTGGACTATCAAGC 486
DB 164486 ctgtgtcttaagactactcttggtgagctctccttaaaagagatgtcacacactgagatt 164545
QY 487 CTGTGTAGCAATTTGTCAATTTATCTTCAGAAATTCCTGTGATTAAGAGATAATT 544
DB 164546 ctgtgttaactgttttggatgtagctcagagatctagctgtccttaaaaaaaaaact 164603

RESULT 6
US-08-998-416-746
Sequence 746, Application US/08998416
Patent No. 6239264
GENERAL INFORMATION:
APPLICANT: Philippsen, Peter
APPLICANT: Pohlmann, Rainer
APPLICANT: Steiner, Sabine
APPLICANT: Mohr, Christine
APPLICANT: Wendland, Jurgen
APPLICANT: Knechtle, Philipp
APPLICANT: Reblischung, Corinne
TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSYPPII
TITLE OF INVENTION: AND USES THEREOF
NUMBER OF SEQUENCES: 1152
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6239264artis Corporation
STREET: 3054 Cornwallis Road
CITY: Research Triangle Park
STATE: No. 6239264th Carolina
COUNTRY: USA
ZIP: 27709
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/998,416
FILING DATE: 24-DEC-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: CH 0016/97
FILING DATE: 31-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8587
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 746:
SEQUENCE CHARACTERISTICS:
LENGTH: 716 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: PAG1484UP
US-08-998-416-746

Query Match 4.7%; Score 35.2; DB 4; Length 716;
Best Local Similarity 52.0%; Pred. No. 0.14;
Matches 79; Conservative 0; Mismatches 73; Indels 0; Gaps 0;
QY 56 GCTCAGGCTTCCAAAGTGGAGCGGACATGTTCTCCCTCAAGAGTGAACGCCGTGG 115
DB 375 GCGGGGGGTGCAACGTGTGTGACAGAGTGTTCACGACAGATGAGCGCCACATCTCG 434

QY 116 CCATGTGAGCTGGAGCTGGAAGCATACGTGCCCATCTGCAAGGCTCCAGGTATG 175
DB 435 ACAGCTGAGCTGGAGCTGACGACGCCGATGGAAGCTACGTCTCAAGCTGAGC 494
QY 176 ATGCTGCTTAGATGTCAGCTGGAACAA 207
DB 495 AGAAGCTTTGATACAGTACGAGCTATACAA 526

RESULT 7

US-09-040-984-58
; Sequence 58, Application US/09040984
; Patent No. 6210883

GENERAL INFORMATION:

APPLICANT: Reed, Steven G.
APPLICANT: Wang, Tongtong
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS
TITLE OF INVENTION: OF LUNG CANCER
NUMBER OF SEQUENCES: 86
CORRESPONDENCE ADDRESS:

ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98104

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/040,984
FILING DATE: 18-MAR-1998

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:
NAME: Makl, David J.

REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.456

TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-622-4900

TELEFAX: 206-282-6031
TELEX:

INFORMATION FOR SEQ ID NO: 58:
SEQUENCE CHARACTERISTICS:

LENGTH: 433 base pairs
TYPE: nucleic acid

STRANDEDNESS: single
TOPOLOGY: linear

US-09-040-984-58

Query Match 4.5%; Score 34.2; DB 4; Length 433;
Best Local Similarity 50.3%; Pred. No. 0.21;

Matches 84; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

QY 538 GATTAATTTATTAAGGTGCTCTCTACCTGTGAGTGTGTGCGACACAGCTTAG 597
DB 103 GCTCATATAGTAAGGATCTCTTCAGTGCAGATGTGAATATGATCATCCAGAGTG 162
QY 598 AAGTGTATATAAAGAGAAAGAGCTCCAAATTTGAATCACTTAATTAATCCATTCTA 657
DB 163 ATGTATCTGTGACAGTCACGCTTTAAAGCTGAACATTTTATGAATACCAATAATA 222
QY 658 TACAACAGCAGTGGAGCACTTTCGAGACTTTTCGATGCTTAGG 704
DB 223 GACCTCTTGACTGAAAACATATTTGTGACTTTAATCGTGTCTTG 269

RESULT 8

US-09-123-912-58
; Sequence 58, Application US/09123912A
; Patent No. 6312695

GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Wang, Tongtong
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY OF LUNG CANCER
FILE REFERENCE: 210121.455C1
CURRENT APPLICATION NUMBER: US/09/123,912A
CURRENT FILING DATE: 1998-07-27
PRIOR APPLICATION NUMBER: 09/040,802
PRIOR FILING DATE: 1998-03-18
NUMBER OF SEQ ID NOS: 114
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 58
LENGTH: 433
TYPE: DNA
ORGANISM: Homo sapiens
US-09-123-912-58

Query Match 4.5%; Score 34.2; DB 4; Length 433;
Best Local Similarity 50.3%; Pred. No. 0.21;

Matches 84; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

QY 538 GATTAATTTATTAAGGTGCTCTCTACCTGTGAGTGTGTGCGACACAGCTTAG 597
DB 103 GCTCATATAGTAAGGATCTCTTCAGTGCAGATGTGAATATGATCATCCAGAGTG 162
QY 598 AAGTGTATATAAAGAGAAAGAGCTCCAAATTTGAATCACTTAATTAATCCATTCTA 657
DB 163 ATGTATCTGTGACAGTCACGCTTTAAAGCTGAACATTTTATGAATACCAATAATA 222
QY 658 TACAACAGCAGTGGAGCACTTTCGAGACTTTTCGATGCTTAGG 704
DB 223 GACCTCTTGACTGAAAACATATTTGTGACTTTAATCGTGTCTTG 269

RESULT 9

US-09-656-286-10/c
; Sequence 10, Application US/09656286
; Patent No. 6355435

GENERAL INFORMATION:

APPLICANT: Wilson, David
APPLICANT: Linz, John E.

APPLICANT: Kaneene, John B.
APPLICANT: Mansfield, Linda S

APPLICANT: Walker, Robert D.
APPLICANT: Newman, Thomas C

TITLE OF INVENTION: METHODS FOR DETECTING AND ENUMERATING CAMPYLOBACTER
TITLE OF INVENTION: JETJUN IN ENVIRONMENTAL SAMPLES AND FOR IDENTIFYING

FILE REFERENCE: MSU 4.1-464/465A
CURRENT APPLICATION NUMBER: US/09/656,286

NUMBER OF SEQ ID NOS: 19
SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 10
LENGTH: 302

TYPE: DNA
ORGANISM: Campylobacter jejuni

US-09-656-286-10

Query Match 4.4%; Score 33.2; DB 4; Length 302;
Best Local Similarity 64.1%; Pred. No. 0.35;

Matches 50; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

QY 227 TCTGGGAGATGTAATCATCTCTCCACACAGCTGCAAGTCCCTGGGTGAACAGA 286
DB 99 TCTTGACCCATTTCAACCAAGACATTAAGCTGTATCTTCATGAGGTATACGA 40

QY 287 ACAATCGTGCCTCTCT 304
DB 39 CCTATACAGCACCCACT 22

RESULT 10
US-09-656-286-12/c
; Sequence 12, Application US/09656286
; Patent No. 6355435
; GENERAL INFORMATION:
; APPLICANT: Wilson, David
; APPLICANT: Linz, John E
; APPLICANT: Kaneene, John B
; APPLICANT: Mansfield, Linda S
; APPLICANT: Walker, Robert D
; APPLICANT: Newman, Thomas C
; TITLE OF INVENTION: METHODS FOR DETECTING AND ENUMERATING CAMPYLOBACTER
; TITLE OF INVENTION: JEJUNI IN ENVIRONMENTAL SAMPLES AND FOR IDENTIFYING
; FILE REFERENCE: MSU 4.1-464/465A
; CURRENT APPLICATION NUMBER: US/09/656,286
; CURRENT FILING DATE: 2000-09-06
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 302
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
US-09-656-286-12

Query Match 4.4%; Score 33.2; DB 4; Length 302;
Best Local Similarity 64.1%; Pred. No. 0.35; Mismatches 28; Indels 0; Gaps 0;
Matches 50; Conservative 0;

QY 227 TCTGGGAGAAATGTAATCTCTCCACAACTGCTGCATGCTCCCTGGGGAAGAAGA 286
DB 99 TCTTGAGCATCTTAACCAAGCATCATTAACCTGCTGATCTCCATGTGGGATACGA 40
QY 287 ACAATGCTGCCCTCTCT 304
DB 39 CCTATAACAGCACCCACT 22

RESULT 11
US-08-470-179-55/c
; Sequence 55, Application US/08470179
; Patent No. 5645994
; GENERAL INFORMATION:
; APPLICANT: Huang Ph.D. Wei Mun
; TITLE OF INVENTION: Method and Compositions for
; TITLE OF INVENTION: Identification of Species in a Sample
; NUMBER OF SEQUENCES: 207
; CORRESPONDENCE ADDRESS:
; ADDRESS: Trask, Britt and Rossa
; STREET: P.O. Box 2550
; CITY: Salt Lake City
; STATE: Utah
; COUNTRY: USA
; ZIP: 84110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/470,179
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sweigert Ph.D. Susan E.
; REGISTRATION NUMBER: 36,289
; REFERENCE/DOCKET NUMBER: 2601
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 801-532-1922
; TELEFAX: 801-531-9168
; INFORMATION FOR SEQ ID NO: 55:

SEQUENCE CHARACTERISTICS:
; LENGTH: 423 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: not relevant
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "gyra gene segment"
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Campylobacter jejuni
US-08-470-179-55

Query Match 4.4%; Score 33.2; DB 1; Length 423;
Best Local Similarity 64.1%; Pred. No. 0.43; Mismatches 28; Indels 0; Gaps 0;
Matches 50; Conservative 0;

QY 227 TCTGGGAGAAATGTAATCTCTCCACAACTGCTGCATGCTCCCTGGGGAAGAAGA 286
DB 170 TCTTGAGCATCTTAACCAAGCATCATTAACCTGCTGATCTCCATGTGGGATACGA 111
QY 287 ACAATGCTGCCCTCTCT 304
DB 110 CCTATAACAGCACCCACT 93

RESULT 12
US-09-371-696-1/c
; Sequence 1, Application US/09371696
; Patent No. 6287777
; GENERAL INFORMATION:
; APPLICANT: Sytkowski, Arthur J.
; APPLICANT: Yang, Meiheng
; TITLE OF INVENTION: NOVEL NPG-1 GENE THAT IS DIFFERENTIALLY EXPRESSED IN PROSTATE
; TITLE OF INVENTION: TUMORS
; FILE REFERENCE: 01948/053002
; CURRENT APPLICATION NUMBER: US/09/371,696
; CURRENT FILING DATE: 1999-08-10
; EARLIER APPLICATION NUMBER: US 09/022,238
; EARLIER FILING DATE: 1998-02-11
; EARLIER APPLICATION NUMBER: US 08/644,326
; EARLIER FILING DATE: 1996-05-10
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1779
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (236)..(1225)
; NAME/KEY: misc_feature
; LOCATION: (1)...(1779)
; OTHER INFORMATION: n = A,T,C or G
US-09-371-696-1

Query Match 4.3%; Score 32.2; DB 4; Length 1779;
Best Local Similarity 50.3%; Pred. No. 2.2; Mismatches 78; Indels 0; Gaps 0;
Matches 79; Conservative 0;

QY 20 GAGAGAAACGCGCCCTCGCTCCACTCCGGGAGCTCAGACGCTCAGTGGGAGGGG 79
DB 995 GGGAGGATGAAGAGCCCTGGGCTGTGCAAGCCGACAGTGCACCTGGCGATGGG 936
QY 80 ACAAGATGTTCTCCCTCAAGAAGTGAACGCGGTGCGCATGTGAGCTGGAGCTGGAGT 139
DB 935 ACCAGGCTTACAGCCGCGGTAGTAGAAGAGACTGGCGGCGGTGGGAGAGAGAGAG 876
QY 140 GCGATACGTGGCCCATCTGCAAGGCTCCAGGTATGGA 176
DB 875 TTATCTCGGTACACGTCCTGCGGAGATGGTGGCGAA 839

	Query Match	4.28;	Score 31.6;	DB 4;	Length 302;
	Best Local Similarity	62.8%;	Pred. No. 1.2;		
	Matches 49;	Conservative	0;	Mismatches 29;	Indels 0;
0y	227	TCGCGGAGACATGTAATCATTCCTTCCACACACGCTGCATGTCCTCCGCGGGAACAACA	286		

Search completed: July 25, 2002, 19:08:16
Job time: 14187 sec

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 25, 2002, 19:10:29 ; Search time 3583.2 Seconds
(without alignments)
4403.496 Million cell updates/sec

Title: US-09-509-779-3

Perfect score: 754
Sequence: 1 ATGGCCGACGTGAGACG.....CAATTAAGTGCAGTTTAA 754

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: GenEmbl:
2: gb.ba:
3: gb.htg:
4: gb.in:
5: gb.ov:
6: gb.pat:
7: gb.pl:
8: gb.pr:
9: gb.ro:
10: gb.sts:
11: gb.sy:
12: gb.un:
13: gb.vl:
14: gb.ba:
15: em.fun:
16: em.hum:
17: em.in:
18: em.in:
19: em.in:
20: em.in:
21: em.in:
22: em.in:
23: em.in:
24: em.in:
25: em.in:
26: em.in:
27: em.in:
28: em.in:
29: em.in:
30: em.in:
31: em.in:
32: em.in:
33: em.in:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result Query Match Length DB ID Description

1	754	100.0	754	9	AF092878	Homo sapi
2	691	91.6	836	9	AF164679	Homo sapi
3	689.4	91.4	816	9	BC008627	Homo sapi
4	663.6	88.0	822	9	BC005966	Homo sapi
5	607.2	80.5	163521	2	AC108040	Homo sapi
6	607.2	80.5	163757	2	AC092418	Homo sapi
7	553.8	73.4	59155	2	AC104983	Homo sapi
8	467.4	62.0	125041	2	AC068693	Homo sapi
9	455.4	60.4	179871	2	AC092978	Homo sapi
10	455.4	60.4	191221	2	AC067952	Homo sapi
11	410.8	54.5	169675	9	AC060766	Homo sapi
12	338.8	44.9	342	9	AF142060	Homo sapi
13	337.2	44.7	342	6	AX282592	Sequence
14	334	44.3	1140	10	AF092877	Homo sapi
15	333.4	44.2	1125	10	BC011127	Mus muscu
16	315.8	41.9	389	6	AX192664	Sequence
17	314.4	41.7	388	6	AX192658	Sequence
18	235	31.2	674	9	AF312226	Homo sapi
19	175.4	23.3	439	9	HS323208	Homo sapi
20	175.4	23.3	191221	2	AC067952	Homo sapi
21	154.6	20.5	160012	2	AC106176	Rattus no
22	154.6	20.5	170087	2	AC098496	Rattus no
23	99.2	13.2	125041	2	AC068693	Homo sapi
24	84.6	11.2	166545	2	AC097192	Rattus no
25	83	11.0	170869	2	AC020266	Drosophill
26	83	11.0	170869	3	AC011696	Drosophill
27	83	11.0	171831	3	AC007473	Drosophill
28	83	11.0	278196	3	AE003825	Drosophill
29	80.4	10.7	219912	2	AC020855	Mus muscu
30	77.8	10.3	5347	6	AX281690	Sequence
31	76.8	10.2	554	9	BC017370	Homo sapi
32	76	10.1	327	9	AF142059	Homo sapi
33	76	10.1	497	9	HUM060A05	Homo sapi
34	76	10.1	508	9	AF140598	Homo sapi
35	76	10.1	544	9	BC001466	Homo sapi
36	75.2	10.0	504	10	AF140599	Mus muscu
37	72.6	9.6	3484	9	HSYEST	X73608 H.sapiens m
38	72.4	9.6	553	3	AF210810	Drosophill
39	72.4	9.6	750	3	AF218290	Drosophill
40	72.4	9.6	3185	3	AY061302	Drosophill
41	72.4	9.6	95294	2	AC019742	Drosophill
42	72.4	9.6	167962	3	AC010010	Drosophill
43	72.4	9.6	310120	3	AE003468	Drosophill
44	70.6	9.4	3208	6	AX212267	Sequence
45	68.8	9.1	10899	3	CEL10A10	U80449 Caenorhabdit

ALIGNMENTS

RESULT 1
AF092878
LOCUS AF092878 754 bp mRNA linear PRI 24-JUL-2001
DEFINITION Homo sapiens zinc RING finger protein SAG mRNA, complete cds.
ACCESSION AF092878
VERSION AF092878.1 GI:4588033
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 (bases 1 to 754)
Swaroop,M., Bian,J., Aviram,M., Duan,H., Blsger,C.L., Loo,J.A.
and Sun,Y.
Expression, purification, and biochemical characterization of SAG,
a RING finger redox-sensitive protein
Free Radical Biol. Med. 27, 193-202 (1999)
2 (bases 1 to 754)
Sun,Y.
Alterations of SAG mRNA in human cancer cell lines: requirement for
the RING finger domain for apoptosis protection
Carcinogenesis. 20 (10), 1899-1903 (1999)
MEDLINE
99435944

PUBMED 10506102
 REFERENCE 3 (bases 1 to 754)
 AUTHORS Duan, H., Wang, Y., Avlram, M., Swaroop, M., Loo, J. A., Bian, J., Tian, Y., Meller, T., Bisgaier, C. L. and Sun, Y.
 TITLE SAG, a novel zinc RING finger protein that protects cells from apoptosis induced by redox agents
 JOURNAL Molecular and cellular biology. 19 (4), 3145-3155 (1999)
 MEDLINE 99182502
 PUBMED 10082581
 REFERENCE 4 (bases 1 to 754)
 AUTHORS Swaroop, M., Wang, Y., Miller, P., Duan, H., Jatkoe, T., Madore, S. J. and Sun, Y.
 TITLE Yeast homolog of human SAG/ROC2/Rbx2/Hrt2 is essential for cell growth, but not for germination: chip profiling implicates its role in cell cycle regulation
 JOURNAL Oncogene. 19 (24), 2855-2866 (2000)
 MEDLINE 20309864
 PUBMED 10851089
 REFERENCE 5 (bases 1 to 754)
 AUTHORS Duan, H., Tsvetkov, L. M., Liu, Y., Song, Y., Swaroop, M., Wen, R., Kung, H. P., Zhang, H. and Sun, Y.
 TITLE Promotion of S-phase entry and cell growth under serum starvation by SAG/ROC2/Rbx2/Hrt2, an E3 ubiquitin ligase component: association with inhibition of p27 accumulation
 JOURNAL Molecular carcinogenesis. 30 (1), 37-46 (2001)
 MEDLINE 21152847
 PUBMED 11255262
 REFERENCE 6 (bases 1 to 754)
 AUTHORS Sun, Y.
 TITLE Direct Submission
 JOURNAL Submitted (16-SEP-1998) Department of Molecular Biology, Parke-Davis, 2800 Plymouth Rd, Ann Arbor, MI 48105, USA
 FEATURES
 source
 1. 754
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="3"
 /map="3q22-q24"
 /cell_line="HeLa, D98/AH-2, HPRt-"
 1. 342
 /function="growth promotion"
 /note="redox sensitive, metal binding; expression protects cells from apoptosis induced by redox compounds"
 /codon_start=1
 /product="zinc RING finger protein SAG"
 /protein_id="AAD25962.1"
 /db_xref="GI:4588034"
 /translation="MADVEDGEETCALASHSGSSGSGDKMFSLKRWNAVAMNSWDVECDTCAICRVQMDACLKQAEKNKOEDCVVWVGECNHSFHNCCMSLWVKNNRCPCLCQDDWVVRIGK"
 BASE COUNT 205 a 155 c 201 g 193 t
 ORIGIN
 Query Match 100.0%; Score 754; DB 9; Length 754;
 Best Local Similarity 100.0%; Pred. No. 2.3e-202;
 Matches 754; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 241 ATCATTCCTTCCACAACACTGCTGATGCTCCCTGGGGTGAACAGAACATTCGTCGCT 300
 |||||||
 Db 241 AATCATTCCTTCCACAACACTGCTGATGCTCCCTGGGGTGAACAGAACATTCGTCGCT 300
 QY 301 CTCTGCCAGCAGGACTGGGTGGTCCAAAGATCGGCAATAGAGTGTGTAAGAGCTTC 360
 |||||||
 Db 301 CTCTGCCAGCAGGACTGGGTGGTCCAAAGATCGGCAATAGAGTGTGTAAGAGCTTC 360
 QY 361 TTAGCGAGTTGTTCAGAGCCCTGGTGATCTGTATCCAGTCCCTACAAAGGCTAGA 420
 |||||||
 Db 361 TTAGCGAGTTGTTCAGAGCCCTGGTGATCTGTATCCAGTCCCTACAAAGGCTAGA 420
 QY 421 ACACACAGGAGGATGATTTCTCAATAGAGGCGATGATCTGTGCTTTGGACTCAT 480
 |||||||
 Db 421 ACACACAGGAGGATGATTTCTCAATAGAGGCGATGATCTGTGCTTTGGACTCAT 480
 QY 481 CAAAGCCTTGTTAGCATTTGTGAGTTTATCTTCAGAAATCTCTGTGATTAAGAGAT 540
 |||||||
 Db 481 CAAAGCCTTGTTAGCATTTGTGAGTTTATCTTCAGAAATCTCTGTGATTAAGAGAT 540
 QY 541 AATTATTAAGGTGGTCCCTTCTACTCTGTGTGTGTGTCGCGACACAGCTTAGAG 600
 |||||||
 Db 541 AATTATTAAGGTGGTCCCTTCTACTCTGTGTGTGTGTCGCGACACAGCTTAGAG 600
 QY 601 TGCTATAAAAAAGAAAGAGCTCCAAATGGAATCACCCTTAATTTACCATTTCTATAC 660
 |||||||
 Db 601 TGCTATAAAAAAGAAAGAGCTCCAAATGGAATCACCCTTAATTTACCATTTCTATAC 660
 QY 661 AACAGCAGTGGGAAGCAGCTTTCAGACTTTTGAAGCTTATGTTGATCAGTTAAAAA 720
 |||||||
 Db 661 AACAGCAGTGGGAAGCAGCTTTCAGACTTTTGAAGCTTATGTTGATCAGTTAAAAA 720
 QY 721 GAATGTACAGTAACAATAATTAAGTGCAGTTTAA 754
 |||||||
 Db 721 GAATGTACAGTAACAATAATTAAGTGCAGTTTAA 754
 RESULT 2
 AF164679 836 bp mRNA linear PRI 14-OCT-1999
 LOCUS AF164679
 DEFINITION Homo sapiens ring finger protein CKBP1 mRNA, complete cds.
 ACCESSION AF164679
 VERSION AF164679.1 GI:5917673
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 836)
 AUTHORS Son, M. Y., Park, J. W., Kim, Y. S., Kang, S. W., Marshak, D. R., Park, W. and Bae, Y. S.
 TITLE Protein kinase CKII interacts with and phosphorylates the SAG
 JOURNAL Biochem. Biophys. Res. Commun. 263 (3), 743-748 (1999)
 MEDLINE 99443734
 REFERENCE 2 (bases 1 to 836)
 AUTHORS Son, M. Y., Park, J. W., Kim, Y. S., Kang, S. W., Marshak, D. R., Park, W. and Bae, Y. S.
 TITLE Direct Submission
 JOURNAL Submitted (01-JUL-1999) Department of Biochemistry, Kyungpook National University, Taegu 702-701, Korea
 FEATURES
 source
 1. 836
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 54. .395
 /note="SAG: contains an H2 ring finger motif"
 /codon_start=1
 /product="ring finger protein CKBP1"
 /protein_id="AAD53984.1"
 /db_xref="GI:5917674"
 /translation="MADVEDGEETCALASHSGSSGSGDKMFSLKRWNAVAMNSWD

VECDTCALCRVQVMDACLRCAENKQEDCVVWVWGCNHSFHNCKMSLWVKONNRCLC
 ODDWVVRIGK*
 BASE COUNT 228 a 185 c 223 g 200 t
 ORIGIN

Query Match 91.6%; Score 691; DB 9; Length 836;
 Best Local Similarity 97.8%; Pred. No. 1.5e-184;
 Matches 744; Conservative 0; Mismatches 10; Indels 7; Gaps 4;

OY 1 ATGGCCGAGCTGGAGAGCAGAGAAACCTGCGCCCTGCTCCTCCTCGGAGACTCA 60
 |||||||
 DB 54 ATGGCCGAGCTGGAGAGCAGAGAAACCTGCGCCCTGCTCCTCCTCGGAGACTCA 113
 OY 61 GGCTCCAACTCGGAGAGCAGAAAGATGTTCTCCCTCAAGAAGTGAACCGGCGGCAATG 120
 |||||||
 DB 114 GGCTCCAACTCGGAGAGCAGAAAGATGTTCTCCCTCAAGAAGTGAACCGGCGGCAATG 173
 OY 121 TGGAGCTGGAGAGCTGGAGATGCGATACGTCGCCATCTGCAGAGGTCCAGGTGATGATGCC 180
 |||||||
 DB 174 TGGAGCTGGAGAGCTGGAGATGCGATACGTCGCCATCTGCAGAGGTCCAGGTGATGATGCC 233
 OY 181 TGTCTTAGATGTCACAGCTGAAACAAACAGAGAGACTGTGTGTGTGGGAGAAATGT 240
 |||||||
 DB 234 TGTCTTAGATGTCACAGCTGAAACAAACAGAGAGACTGTGTGTGTGGGAGAAATGT 293
 OY 241 AATCATTTCTTCACAACTGCTGCATGTCCTGCTGGTGAACAGAAACATGCTCCCT 300
 |||||||
 DB 294 AATCATTTCTTCACAACTGCTGCATGTCCTGCTGGTGAACAGAAACATGCTCCCT 353
 OY 301 CTCTGCGACAGAGACTGGGTGTCCAAAGATCGCAATGAGAGTGTAGAGAGCTTC 360
 |||||||
 DB 354 CTCTGCGACAGAGACTGGGTGTCCAAAGATCGCAATGAGAGTGTAGAGAGCTTC 413
 OY 361 TTAGGCACTTTGTCAGAGCCCTGCTGATCTTTGATTCAGAGTCCCTACAAAGCTTGA 420
 |||||||
 DB 414 TTAGGCACTTTGTCAGAGCCCTGCTGATCTTTGATTCAGAGTCCCTACAAAGCTTGA 473
 OY 421 ACACATCAGAGGATGATCTTCAAAATAGAGCCGATGATGCTGTC--TTTGGAGTC 478
 |||||||
 DB 474 ACACATCAGAGGATGATCTTCAAAATAGAGCCGATGATGCTGTC--TTTGGAGTC 533
 OY 479 ATCAAAAGCCTTGTT--AGCATTTTGTCACTTTATCTTCAGAAATCTCTGTGATTAGA 536
 |||||||
 DB 534 ATCAAAAGCCTTGTT--AGCATTTTGTCACTTTATCTTCAGAAATCTCTGTGATTAGA 593
 OY 537 AATATATTTATTAAGAGTGTCTCTCTACCTCTGTGTGTGTGTGCGCACACAGCTTA 596
 |||||||
 DB 594 AATATATTTATTAAGAGTGTCTCTCTACCTCTGTGTGTGTGTGCGCACACAGCTTA 653
 OY 597 GAAGTGTCTTAAAGAAAGAGAGCTCAAAATTTGAATCAGC--TTTAAATTTACCATTTTC 655
 |||||||
 DB 654 GAAGTGTCTTAAAGAAAGAGAGAGCTCAAAATTTGAATCAGC--TTTAAATTTACCATTTTC 713
 OY 656 TATACACAGCAGTGAAGAGAGCTTC--GAGACTTTTTCAGTCTTATGTTGATGACT 713
 |||||||
 DB 714 TATACACAGCAGTGAAGAGAGCTTC--GAGACTTTTTCAGTCTTATGTTGATGACT 773
 OY 714 TAAAAAAGATGTTACAGTAACAAATTAAGTCACTTTAAA 754
 |||||||
 DB 774 TAAAAAAGATGTTACAGTAACAAATTAAGTCACTTTAAA 814

RESULT 3
 BC008627 816 bp mRNA linear PRI 12-JUL-2001
 LOCUS Homo sapiens, ring finger protein 7, clone MGC:117274 IMAGE:4177613,
 DEFINITION mRNA, complete cds.
 ACCESSION BC008627
 VERSION BC008627.1 GI:14250388
 KEYWORDS MGC.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE

1 (bases 1 to 816)
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 Strausberg, R.
 Direct Submission
 Submitted (25-MAY-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA

REMARK

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: David N. Louis, M.D.
 cdna Library Preparation: Life Technologies, Inc.
 cdna Library Arrayed by: The I.M.A.G.E. Consortium (LINL)
 DNA Sequencing by: Baylor College of Medicine Human Genome
 Sequencing Center
 Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: Villalón, D.K., Luna, R.A., Hale, S.M., Huylk, S., Lu, X., Garcia,
 A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W.,
 Muzny, D.M., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LINL at: <http://image.llnl.gov>
 Series: IRK Plate: 12 Row: h Column: 14
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Similarity but not
 identity to protein.

FEATURES

Location/Qualifiers
 1..816
 /organism="Homo sapiens"
 /db_xref="LOCUSID:9616"
 /db_xref="taxon:9606"
 /clone="MGC:117274 IMAGE:4177613"
 /tissue_type="Brain, anaplastic oligodendroglioma with
 1p/19q loss"
 /clone_lib="NCI CGAP_Brn67"
 /lab_host="DH10B"
 /note="Vector: PCMV-SPORT6"
 37..378
 /codon_start=1
 /product="ring finger protein 7"
 /protein_id="AAH08627.1"
 /db_xref="GI:14250388"
 /translation="MADVEDERETCALASHSGSGSKSGDKMFSLKRMNAVAMMSMD
 VECDTCALCRVQVMDACLRCAENKQEDCVVWVWGCNHSFHNCKMSLWVKONNRCLC
 ODDWVVRIGK"
 ODDWVVRIGK*

CDS

BASE COUNT 226 a 178 c 213 g 199 t
 ORIGIN

Query Match 91.4%; Score 689.4; DB 9; Length 816;
 Best Local Similarity 97.6%; Pred. No. 4.5e-184;
 Matches 743; Conservative 0; Mismatches 11; Indels 7; Gaps 4;

OY 1 ATGGCCGAGCTGGAGAGCAGAGAAACCTGCGCCCTGCTCCTCCTCGGAGACTCA 60
 |||||||
 DB 37 ATGGCCGAGCTGGAGAGCAGAGAAACCTGCGCCCTGCTCCTCCTCGGAGACTCA 96
 OY 61 GGCTCCAACTCGGAGAGCAGAAAGATGTTCTCCCTCAAGAAGTGAACCGGCGGCAATG 120
 |||||||
 DB 97 GGCTCCAACTCGGAGAGCAGAAAGATGTTCTCCCTCAAGAAGTGAACCGGCGGCAATG 156
 OY 121 TGGAGCTGGAGAGCTGGAGATGCGATACGTCGCCATCTGCAGAGGTCCAGGTGATGATGCC 180
 |||||||
 DB 157 TGGAGCTGGAGAGCTGGAGATGCGATACGTCGCCATCTGCAGAGGTCCAGGTGATGATGCC 216
 OY 181 TGTCTTAGATGTCACAGCTGAAACAAACAGAGAGACTGTGTGTGTGTGTGGGAGAAATGT 240
 |||||||
 DB 217 TGTCTTAGATGTCACAGCTGAAACAAACAGAGAGACTGTGTGTGTGTGTGGGAGAAATGT 276

OY	241	AATCATTTCTTCCACAAAGTCGTGCAATGTCCTCGTGGGTGAATAACAGAACAAATGCCTCCT	300
OY	277	AATCATTTCTTCCACAAAGTCGTGCAATGTCCTCGTGGGTGAATAACAGAACAAATGCCTCCT	336
OY	301	CTCTGCCACAGAGACTGGGTGGTCCAAAAGATGGCAAATGAGAGTGGTTAGAAGGCTTC	360
Db	337	CTCTGCCACAGAGACTGGGTGGTCCAAAAGATGGCAAATGAGAGTGGTTAGAAGGCTTC	396
OY	361	TTAGGCGCATTTGTTACAGACCCTGGTGGATCTTTGTAATCATCACTGGCCCTACAAAGGCTAGA	420
Db	397	TTAGGCGCATTTGTTACAGACCCTGGTGGATCTTTGTAATCATCACTGGCCCTACAAAGGCTAGA	456
OY	421	ACACTACAGAGGATGAATTCTCAATATAGAGCCCATGATGATGTGGTC - TTTGGAGTC	478
Db	457	ACACTACAGAGGATGAATTCTCAATATAGAGCCCATGATGATGTGGTC - TTTGGAGTC	516
OY	479	ATCAAAAGCCCTTGTT - AGCATTTGTCACTTTTATCTTCAGAAATCTCTGTGATTAAGA	536
Db	517	ATCAAAAGCCCTTGTTAGCATTTTGTGCAAGTTTATCTTCAGAAATCTCTGTGATTAAGA	576
OY	537	AGATATATTTATTAAGGTGGTCTCTTCCTCAACCTCTGTGTGTGTGTCGGCACACAGCTTA	596
Db	577	AGATATATTTATTAAGGTGGTCTCTTCCTCAACCTCTGTGTGTGTGTCGGCACACAGCTTA	636
OY	597	GAAAGTGCTATTAATAAAGGAAGAGCTCCAATTTGAATGAC - TTATTAATTTCCCATTTTC	655
Db	637	GAAAGTGCTATTAATAAAGGAAGAGCTCCAATTTGAATGAC - TTATTAATTTCCCATTTTC	696
OY	655	TATACACAGCAGCATGTGAAGCAGTTTC - GAGACTTTTTCAGATGCTTATGTTGATCAGT	713
Db	697	TATACACAGCAGCATGTGAAGCAGTTTCAGAGAACTTTTGCATGCTTATGTTGATCAGT	756
OY	714	TAAAAAGAGTGTACAGTACACAATTAAGTCAGTTTTAA	754
Db	757	TAAAAAGAGTGTACAGTACACAATTAAGTCAGTTTTAA	797
RESULT	4		
LOCUS	BC005966	822 bp mRNA linear	PRI 12-JUL-2001
DEFINITION	Homo sapiens, ring finger protein 7, clone MGC:14618 IMAGE:4069078,		
ACCESSION	BC005966	mRNA, complete cds.	
VERSION	BC005966.1	GI:13543635	
KEYWORDS	MGC.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	1 (bases 1 to 822)		
JOURNAL	Strausberg, R. Direct Submission Submitted (02-APR-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA		
REMARK	NIH-MGC Project URL: http://mgc.ncl.nih.gov		
COMMENT	Contact: MGC help desk Email: cgapbs-r@mail.nih.gov Tissue Procurement: ATCC CDNA Library Preparation: CLONTECH Laboratories, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNLI) DNA Sequencing by: Sequencing Group at the Stanford Human Center, Stanford University School of Medicine, Stanford, CA 94305 Web site: http://www.sbcg.stanford.edu Contact: (Dickson, Mark) mcdex@xll.stanford.edu Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.		

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: http://image.llnl.gov Series: IRAL Plate: 21 Row: o Column: 13 This clone was selected for full length sequencing because it

FEATURES					
	Source	Location/Qualifiers	passed the following selection criteria: Genomescan gene prediction, Similarity but not identity to protein.		
		1..822	/organism="Homo sapiens"		
			/db_xref="locusID:9616"		
			/db_xref="taxon:9606"		
			/clone="MGC:14618 IMAGE:4069078"		
			/tissue_type="Kidney, hypernephroma"		
			/clone_id="NH_MGC_58"		
			/lab_host="DH10B"		
			/note="Vector: pDNR-LIB"		
CDS		27..368	/codon_start=1		
			/product="ring finger protein 7"		
			/protein_id="AAH05966.1"		
			/db_xref="GI:13543636"		
			/translation="MADVEDDEETKALASHIGSSGSKSGDKMFSLKKINAVAMSWMDVECTCAICRYOVMADCLRCQAEKQEDCVVWMECHNPSFINCCMSLWVKNNNCPLCQQDWYVORIK"		
BASE COUNT	240 a	175 c	210 g	197 t	
ORIGIN					
Query Match	88.0%	Score 663.6;	DB 9;	Length 822;	
Best Local Similarity	97.1%;	Pred. No. 9e-177;	Mismatches 14;	Indels 8;	Gaps 6;
Matches 740;	Conservative	0;	Partial matches		
Oy	1	ATGGCGAGCTGTGAAGAGAGAGAAMAACCTGCGCCTTCACCTCCAGTCCGAGACTCA	60		
Db	27	ATGGCGAGCTGTGAAGAGAGAGAAMAACCTGCGCCTTCACCTCCAGTCCGAGACTCA	86		
Oy	61	GGCTCCAAGTCGGGAGGCGACAAGATTCTCCCTCAAGAAGTGAACGCGTGCCATG	120		
Db	87	GGCTCCAAGTCGGGAGGCGACAAGATTCTCCCTCAAGAAGTGAACGCGTGCCATG	146		
Oy	121	TGGAGCTGGGACGTGGAGTGGCATACGTGCCCATCTCGAGGCTCAGATGATGATGCC	180		
Db	147	TGGAGCTGGGACGTGGAGTGGCATACGTGCCCATCTCGAGGCTCAGATGATGATGCC	206		
Oy	181	TGCTTAGATGTCAAGCTGTGAANAACAACAGAGACTGTGTTGTGCTGGGGAGATGT	240		
Db	207	TGCTTAGATGTCAAGCTGTGAANAACAACAGAGACTGTGTTGTGCTGGGGAGATGT	266		
Oy	241	AATCATTCCTTCACACACATGCTGCATGTCCTGTGGGTGAACAAGAACATGCTGCCCT	300		
Db	267	AATCATTCCTTCACACATGCTGCATGTCCTGTGGGTGAACAAGAACATGCTGCCCT	326		
Oy	301	CTCTGCCAGCAGAGACTGGGTGGTCCAAAGATTCGCAATAGAGTGGTTAGAAAGCTTC	360		
Db	327	CTCTGCCAGCAGAGACTGGGTGGTCCAAAGATTCGCAATAGAGTGGTTAGAAAGCTTC	386		
Oy	361	TTAGGCGAGTGTTCAGAGCCCTGGTGGATCTTGTAAATCCAGTGGCTACAAAGGCTTGA	420		
Db	387	TTAGGCGAGTGTTCAGAGCCCTGGTGGATCTTGTAAATCCAGTGGCTACAAAGGCTTGA	446		
Oy	421	ACACTACAGGGGATGAATTCCTCAATATAGAGACCAGATGATCTGTGATC - "TTTTGACTC	478		
Db	447	ACACTACAGGGGATGAATTCCTCAATATAGAGACCAGATGATCTGTGATC "TTTTGACTC	506		
Oy	479	ATCAAAGCCCTTGG - TTAGACTTTTGTCACTTTATCTTAGAAAATTCCTGTGATTAAGAA	537		
Db	507	ATCAAAGCCCTTGGTTTACCATTTTGTCTAGTTATCTTAGAAAATTCCTGTGATTAAGAA	566		
Oy	538	GATAATTTAATAAGGTCCTTCCTCACTGCTGCTGGTGTGTCGCGCACACAGCTTAG	597		
Db	567	GATAATTTAATAAGGTCCTTCCTCACTGCTGCTGGTGTGTCGCGCACACAGCTTAG	626		
Oy	598	AAGTCTCAT - AAAAAAGAAAGAGCTCC - AAATTGAATCACC - TTAATATTACCATTT	654		
Db	627	AAGTCTCAT AAAAAAGAAAGAGCTCCAAATATGAATCACCCTTATATATTACCATTT	686		
Oy	655	CTATPACAAGCAGAGTGAACAGTTTC - GAGACTTTTTCGATGCTATGATGATATCAG	712		

```
Db 687 CTATACACAGCAGTGGAGAGCTTTTCAGAGAACTTTTGCATGCTTATGGTTGATCAG 746
QY 713 TTTAAAAAGATGTTACATACAAATTAAGTCAGTTTAA 754
Db 747 TTTAAAAAGATGTTACATACAAATTAAGTCAGTTTAA 788

RESULT 5
AC108040/c 163521 bp DNA linear HTG 30-JAN-2002
LOCUS Homo sapiens chromosome 4 clone RP11-210010, WORKING DRAFT
DEFINITION AC108040
SEQUENCE 4 unordered pieces.
AC108040
AC108040.2 GI:18425316
HTG: HTGS_PHASE1, HTGS_DRAFT; HTGS_FULLTOP.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 163521)
Waterston, R.H.
The sequence of Homo sapiens clone
2 (bases 1 to 163521)
Waterston, R.H.
Unpublished
Reference
AUTHORS Submitted (24-JAN-2002) Genome Sequencing Center, Washington
TITLE University School of Medicine, 4444 Forest Park Parkway, St. Louis,
JOURNAL MO 63108, USA
COMMENT On Jan 30, 2002 this sequence version replaced gi:18308830.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
Contact: submissions@watson.wustl.edu
Project Information
Center project name: H.NH0210010
----- Summary Statistics -----
Sequencing vector: M13, 0%
Sequencing vector: plasmid, 100%
Chemistry: Dye-primer ET; 0% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 162304 bases at least Q40
Consensus quality: 162444 bases at least Q30
Consensus quality: 162529 bases at least Q20
Insert size: 189000; agarose-fp
Insert size: 163221; sum-of-contigs
Quality coverage: 8.88 in Q20 bases; agarose-fp
Quality coverage: 8.85 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 1564: contig of 1564 bp in length
* 1565 1664: gap of unknown length
* 1665 88984: contig of 87320 bp in length
* 88985 89084: gap of unknown length
* 89085 112702: contig of 23618 bp in length
* 112703 112802: gap of unknown length
* 112803 163521: contig of 50719 bp in length.
*
Location/Qualifiers
1..163521
/organism="Homo sapiens"
/db_xref="taxon:9606"

/chromosome="4"
/clone="RP11-210010"
1..1564
/note="assembly_name:Contig25"
vector_end:SP6
clone_end:left"
1665..88984
/note="assembly_name:Contig28"
clone_end:T7
vector_side:left"
89085..112702
/note="assembly_name:Contig26"
112803..163521
/note="assembly_name:Contig27"
BASE COUNT 51049 a 32426 c 33346 g 46399 t 301 others
ORIGIN

Query Match 80.5%; Score 607.2; DB 2; Length 163521;
Best Local Similarity 90.6%; Pred. No. 3.2e-160;
Matches 696; Conservative 0; Mismatches 58; Indels 14; Gaps 4;

QY 1 ATGGCCGACGTGGAGAGCAGAGAAACCTGGCCCTCTCTACTCCGGAGCTCA 60
Db 80418 ATGGCCGACGTGGAGAGCAGAGAAACCTACGCCCTGACCTCAGAGCGCA 80359
QY 61 GGCTCCAGTGGGAGGCGCAAAAGTCTCCCTCAAGAGTGAAGCGGCGCATG 120
Db 80358 GGCTCCAGTGGGAGGCGCAAAAGTCTCCCTCAAGAGTGAAGCGGCGCATG 80299
QY 121 TGGAGCTGGAGCTGGAGTGCATACGTGCGCATCTGCAGGTCAGGTGATGATGCC 180
Db 80298 TGGAGCTGGAGCTGGAGTGCATACGTGCGCATCTGCAGGTCAGGTGATGATGCC 80239
QY 181 TGTCTTAATGATGCACGCTAAACAAAGAGAGCTGTGTGTGCTGGGAGATGT 240
Db 80238 TGTCTTAATGATGCACGCTAAACAAAGAGAGCTGTGTGTGCTGGGAGATGT 80179
QY 241 AATCATTCCTTCACAAAGCTGCATGCTCCGTGGGTGAACAAATGCTGCCCT 300
Db 80178 AATCATTCCTTCACAAAGCTGCATGCTCCGTGGGTGAACAAATGCTGCCCT 80119
QY 301 CTCTGCCACAGACAGCTGGGTGTCCTCAAAAGATCGCAATGAGAGTGTAGAGCTTC 360
Db 80118 CTCTGCCACAGACAGCTGGGTGTCCTCAAAAGATCGCAATGAGAGTGTAGAGCTTC 80059
QY 361 TTAGCGCAGTGTTCAGAGCCCTGTGTGATCTTGTATCCAGTCCCTCAAAAGCTTAA 420
Db 80058 TTAGCGCAGTGTTCAGAGCCCTGTGTGATCTTGTATCCAGTCCCTCAAAAGCTTAA 79999
QY 421 ACACCTACAGGGGATGATCTTCAAATAGAGCGCATGTGNGTC--TTTGGACTC 478
Db 79998 ACACCTACAGGGGATGATCTTCAAATAGAGCGCATGTGNGTC--TTTGGACTC 79939
QY 479 ATCAAAAGCCTGGTT--AGCATTTGTCACTTTATCTTCAAGAAATCTCTGTGATTAAGA 536
Db 79938 ATCAAAAGCCTGGTT--AGCATTTGTCACTTTATCTTCAAGAAATCTCTGTGATTAAGA 79879
QY 537 AGATTAATTTATTAAGAGTGTCTTCTTACCTGTGTGTGTGTGTGCGCACACACTTA 596
Db 79878 AGATTAATTTATTAAGAGTGTCTTCTTACCTGTGTGTGTGTGTGCGCACACACTTA 79819
QY 597 GAAGTCTATTAATAAAGAAAGAGCGCAATTTGATACAC--TTATTAATTTCCACTTTC 655
Db 79818 GAAGTCTATTAATAAAGAGAGAGACTCCAAACTGAAATACCTTTATTAATTTCCACTTTC 79759
QY 656 TTATCAACAGCGAGTGAAGCAAGTTCT-----GAGACTTTTTCAGTCTTATGTTT 706
Db 79758 TTATCAACAGCGAGTGAAGCAAGTTCT-----GAGACTTTTTCAGTCTTATGTTT 79699
QY 707 GATCACTTAATAAAGAAATGTACAGTACAAATTAAGTCAGTTTAA 754
Db 79698 GATCACTTAATAAAGAAATGTACAGTACAAATTAAGTCAGTTTAA 79651
```

```
RESULT 6
AC092418/c 163757 bp DNA linear HTG 04-JUL-2001
LOCUS
DEFINITION Homo sapiens chromosome 3 clone RP11-229A12, WORKING DRAFT
ACCESSION AC092418 AC016923
VERSION AC092418.1 GI:14595799
KEYWORDS HTGS_PHASE1, HTGS_DRAFT, HTGS_FULLTOP, HTGS_ACTIVEFIN.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 163757)
Kaul, R.K., Olson, M.V., Raymond, C., Clendinning, J., Ivey, R.G. and
Haugen, E.D.
DIRECT SUBMISSION
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 163757)
AUTHORS Kaul, R.K., Olson, M.V., Raymond, C., Clendinning, J., Ivey, R.G. and
Haugen, E.D.
TITLE Direct Submission
JOURNAL Submitted (04-JUL-2001) Genome Center, University of Washington,
REFERENCE Box 352145, Seattle, WA 98195, USA
AUTHORS On Jul 4, 2001 this sequence version replaced gi:8705848.
COMMENT ----- Genome Center
Center: University of Washington Genome Center
Center Code: UWGC
Web site: http://www.genome.washington.edu
Contact: uwgchgs@u.washington.edu
Drafting Center: BCM
----- Project Information
Center clone name: RP11-229A12 (bc0315)
----- Summary Statistics
Sequencing vector: plasmid; 50% of reads
Sequencing vector: unknown; 50% of reads
Chemistry: Dye-terminator ET; 47% of reads
Chemistry: Dye-terminator Big Dye; 52% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 162824 bases at least Q40
Consensus quality: 163281 bases at least Q30
Consensus quality: 163468 bases at least Q20
Insert size: 163557; sum-of-ctrls
Quality coverage: 7.6x in Q20 bases; sum-of-ctrls
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 15071: contig of 15071 bp in length
* 15072 15171: gap of unknown length
* 15172 57886: contig of 42715 bp in length
* 57887 57987: gap of unknown length
* 57987 163757: contig of 105771 bp in length.
* Location/Qualifiers
1. 163757
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="3"
/clone="RP11-229A12"
/clone_lib="RPCT Human BAC library 11"
1. 15071
/misc_feature /note="assembly_name:Contig6"
misc_feature 15172..57886
/misc_feature /note="assembly_name:Contig7"
57987..163757
/misc_feature /note="assembly_name:Contig8"
```


* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

1
725 824: gap of 100 bp in length
825 1582: contig of 758 bp in length
1583 1682: gap of 100 bp
1683 2409: contig of 727 bp in length
2410 2509: gap of 100 bp
2510 3235: contig of 726 bp in length
3236 3335: gap of 100 bp
3336 4091: contig of 755 bp in length
4091 4190: gap of 100 bp
4191 4955: contig of 765 bp in length
4956 5055: gap of 100 bp
5056 5780: contig of 725 bp in length
5781 5880: gap of 100 bp
5881 6643: contig of 763 bp in length
6644 6743: gap of 100 bp
6744 7496: contig of 753 bp in length
7497 7596: gap of 100 bp
7597 8348: contig of 752 bp in length
8349 8448: gap of 100 bp
8449 9217: contig of 769 bp in length
9218 9317: gap of 100 bp
9318 10093: contig of 776 bp in length
10094 10193: gap of 100 bp
10194 10924: contig of 731 bp in length
10925 11024: gap of 100 bp
11025 11793: contig of 769 bp in length
11794 11893: gap of 100 bp
11894 12661: contig of 768 bp in length
12662 12761: gap of 100 bp
12762 13525: contig of 764 bp in length
13526 13625: gap of 100 bp
13626 14347: contig of 722 bp in length
14348 14447: gap of 100 bp
14448 15172: contig of 725 bp in length
15173 15272: gap of 100 bp
15273 16007: contig of 735 bp in length
16008 16107: gap of 100 bp
16108 16855: contig of 748 bp in length
16856 16955: gap of 100 bp
16956 17711: contig of 756 bp in length
17712 17811: gap of 100 bp
17812 18568: contig of 757 bp in length
18569 18668: gap of 100 bp
18669 19421: contig of 753 bp in length
19422 19521: gap of 100 bp
19522 20229: contig of 708 bp in length
20230 20329: gap of 100 bp
20330 21061: contig of 732 bp in length
21062 21161: gap of 100 bp
21162 21909: contig of 748 bp in length
21910 22009: gap of 100 bp
22010 22773: contig of 764 bp in length
22774 22873: gap of 100 bp
22874 23623: contig of 750 bp in length
23624 23723: gap of 100 bp
23724 24489: contig of 766 bp in length
24490 24589: gap of 100 bp
24590 25354: contig of 765 bp in length
25355 25454: gap of 100 bp
25455 26174: contig of 720 bp in length
26175 26274: gap of 100 bp
26275 27043: contig of 769 bp in length
27044 27143: gap of 100 bp
27144 27900: contig of 757 bp in length
27901 28000: gap of 100 bp
28001 28753: contig of 753 bp in length
28754 28853: gap of 100 bp
28854 29572: contig of 719 bp in length

* 29573 29672: gap of 100 bp
29673 30419: contig of 747 bp in length
30420 30519: gap of 100 bp
30520 31262: contig of 743 bp in length
31263 31362: gap of 100 bp
31363 32052: contig of 690 bp in length
32053 32152: gap of 100 bp
32153 32909: contig of 757 bp in length
32910 33009: gap of 100 bp
33010 33736: contig of 727 bp in length
33737 33836: gap of 100 bp
33837 34598: contig of 762 bp in length
34599 34698: gap of 100 bp
34699 35456: contig of 758 bp in length
35457 35556: gap of 100 bp
35557 36274: contig of 718 bp in length
36275 36374: gap of 100 bp
36375 37111: contig of 737 bp in length
37112 37211: gap of 100 bp
37212 37954: contig of 743 bp in length
37955 38054: gap of 100 bp
38055 38783: contig of 729 bp in length
38784 38883: gap of 100 bp
38884 39686: contig of 803 bp in length
39687 39786: gap of 100 bp
39787 40535: contig of 749 bp in length
40536 40635: gap of 100 bp
40636 41387: contig of 752 bp in length
41388 41487: gap of 100 bp
41488 42254: contig of 767 bp in length
42255 42354: gap of 100 bp
42355 43046: contig of 692 bp in length
43047 43146: gap of 100 bp
43147 43885: contig of 739 bp in length
43886 43985: gap of 100 bp
43986 44741: contig of 756 bp in length
44742 44841: gap of 100 bp
44842 45604: contig of 763 bp in length
45605 45704: gap of 100 bp
45705 46429: contig of 725 bp in length
46430 46529: gap of 100 bp
46530 47244: contig of 715 bp in length
47245 47344: gap of 100 bp
47345 48088: contig of 744 bp in length
48089 48188: gap of 100 bp
48189 48948: contig of 760 bp in length
48949 49048: gap of 100 bp
49049 49803: contig of 755 bp in length
49804 49903: gap of 100 bp
49904 50648: contig of 745 bp in length
50649 50748: gap of 100 bp
50749 51506: contig of 758 bp in length
51507 51606: gap of 100 bp
51607 52345: contig of 739 bp in length
52346 52445: gap of 100 bp
52446 53185: contig of 740 bp in length
53186 53285: gap of 100 bp
53286 54059: contig of 774 bp in length
54060 54159: gap of 100 bp
54160 54901: contig of 742 bp in length
54902 55001: gap of 100 bp
55002 55740: contig of 739 bp in length
55741 55840: gap of 100 bp
55841 56577: contig of 737 bp in length
56578 56677: gap of 100 bp
56678 57433: contig of 756 bp in length
57434 57533: gap of 100 bp
57534 58290: contig of 757 bp in length

Query Match 73.4% Score 553.8 DB 2: Length 59155;
Best Local Similarity 89.9% Pred. No. 3.2e-145;
Matches 642; Conservative 0; Mismatches 58; Indels 14; Gaps 4;


```

OY 45 TCACCTCCGGAGCTCAGGCTCCCAAGTCGGAGGCGCAAGATGTTCTCCCTCAAGACATG 104
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 7635 TCACCTCCAGAGCGGAGGCTCCAGTCGGAGGCGCAAGATGTTCTTACTTCAAGAGTG 7694
OY 105 GAACGGCGTCCAGTATGATCATCTCTCCACAACTGCTCATGCTGCTGCTGCTGCTGCT 164
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 7695 GAACGGCGTCCAGTATGATCATCTCTCCACAACTGCTCATGCTGCTGCTGCTGCTGCT 7754
OY 165 CCAGGTGAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 224
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 7755 CCAGGTGAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7814
OY 225 GGTCTGGAGAGATGATGATCATCTCTCCACAACTGCTCATGCTGCTGCTGCTGCTGCT 284
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 7815 GGTCTGGAGAGATGATGATCATCTCTCCACAACTGCTCATGCTGCTGCTGCTGCTGCT 7874
OY 285 GAACAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 344
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 7875 GAACAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7934
OY 345 GTGGTGAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 404
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 7935 GTGGTGAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7994
OY 405 CCTTCAAAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 464
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 7995 CCTTCAAAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8054
OY 465 TGGTC-TTGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 520
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 8055 TGGTC-TTGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8114
OY 521 TTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 580
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 8115 TTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8174
OY 581 TGGCCACACAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 639
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 8175 TGGCCACACAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8234
OY 640 ATAATTTACCACTTCTATACAAAGGAGTGAAGCAGTTTC-----GAGACTTT 690
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 8235 ATAATTTACCACTTCTATACAAAGGAGTGAAGCAGTTTCATTCATTCAGAGAGAACTTT 8294
OY 691 TTGGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 744
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 8295 TTGGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8348

```

```

RESULT 8
AC068693/C 125041 bp DNA linear HTG 04-MAY-2001
LOCUS Homo sapiens chromosome 17 clone RP11-144C9 map 17, WORKING DRAFT
DEFINITION AC068693
SEQUENCE, 25 unordered pieces.
ACCESSION AC068693
VERSION AC068693.2 GI:8139716
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Eumetazoa; Chordata; Chordata; Chordata; Chordata; Chordata; Chordata;
Mammalia; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 125041)
AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Abram, H., Allen, N.,
TITLE Homo sapiens chromosome 17, clone RP11-144C9
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 125041)

```

Birren, B., Linton, L., Nusbaum, C., Lander, E., Abram, H., Allen, N.,
 Anderson, S., Baldwin, J., Barna, N., Bastien, V., Bedalov, F.,
 Bonafant, L., Bouckgeer, B., Brown, A., Burdett, G.,
 Campasiano, A., Castle, A., Choquet, Y., Colangelo, M., Collins, S.,
 Collymore, A., Cooke, P., DeRubeis, K., Dewar, K., Diaz, J., S.,
 Dodge, S., Domingo, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D.,
 Galagan, J., Gardy, S., Glendon, S., Goyette, M., Graham, L.,

TITLE JOURNAL COMMENT

Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L.,
 Howland, J. C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,
 Klein, J., Labrecque, K., Lamazares, R., Landers, T., Lehotzky, J.,
 Levine, R., Liu, C., Liu, G., Locke, K., MacDonald, P., Margulis, N.,
 McCarthy, M., McEwan, P., McGovern, A., McKernan, K., McSheeters, R.,
 Meldrum, J., Meneus, L., Mihova, T., Miranda, C., Miñana, V., Morrow, J.,
 Murphy, T., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P.,
 O'Neill, D., Oliver, J., Oliver, J., Peterson, K., Pierre, N.,
 Pisanic, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,
 Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,
 Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
 Testa, S., Theodore, J., Threlkell, A., Travers, M., Triggillo, J.,
 Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J.,
 Young, G., Zahoun, J., Zimmer, A. and Zody, M.

Submitted (06-MAY-2000) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Jun 1, 2000 this sequence version replaced gi:171215.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/MW/repeatmasker.html

----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Web site: http://www-seq.wi.mit.edu
 Center code: WIBR
 Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information
 Center project name: 144_C-9
 Center clone name: 18985

----- Summary Statistics
 Sequencing vector: M13: M77815: 100% of reads
 Chemistry: Dye-terminator Big Dye: 100% of reads
 Assembly program: Phrap: version 0.960731
 Consensus quality: 112098 bases at least Q40
 Consensus quality: 118849 bases at least Q30
 Consensus quality: 121119 bases at least Q20
 Insert size: 158000; agarose-fp
 Insert size: 122641; sum-of-contigs
 Quality coverage: 3.8 in Q20 bases; agarose-fp
 Quality coverage: 4.9 in Q20 bases; sum-of-contigs

----- NOTE: This is a 'working draft' sequence. It currently
 consists of 25 contigs. The true order of the pieces
 is not known and their order in this sequence record is
 arbitrary. Gaps between the contigs are represented as
 runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 as soon as it is available and the accession number will
 be preserved.

```

1 1019: contig of 1019 bp in length
* 1020 1119: gap of 100 bp
* 1120 3013: contig of 1894 bp in length
* 3014 3113: gap of 100 bp
* 3114 5008: contig of 1895 bp in length
* 5009 5108: gap of 100 bp
* 5109 7154: contig of 2046 bp in length
* 7155 7254: gap of 100 bp
* 7255 10405: contig of 3151 bp in length
* 10406 10505: gap of 100 bp
* 10506 12914: contig of 2409 bp in length
* 12915 13014: gap of 100 bp
* 13015 13781: contig of 767 bp in length
* 13782 13881: gap of 100 bp
* 13882 17746: contig of 3865 bp in length
* 17747 17846: gap of 100 bp
* 17847 20494: contig of 2648 bp in length
* 20495 22173: gap of 100 bp
* 22174 22273: contig of 100 bp in length
* 22274 25788: contig of 3515 bp in length
* 25789 25888: gap of 100 bp
* 25889 30218: contig of 4330 bp in length
* 30219 30318: gap of 100 bp

```


Alsbrooks,S.L., Amaralunge,H.C., Are,J.R., Banks,T., Barbara,J.,
 Benton,J., Bimega,K., Blankenburg,K., Bonnin,D., Bouck,J.,
 Bowles,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buby,C.,
 Burck,P., Burkett,C., Burrell,K.L., Byrd,N.C., Caron,T.F.,
 Carter,N., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
 Chen,Z., Chowdhury,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
 Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
 Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
 Dem,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,
 Dugan-Recha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
 Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,
 Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
 Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
 Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A.,
 Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C.,
 Hollins,B., Homs,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,
 Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
 Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
 Kovar,C., Kratochvil,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
 Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,C., Liu,W.,
 Louisedge,H., Lozada,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
 Ma,J., Maheshwari,M., Kapua,P., Martin,R., Martindale,A.,
 Martinez,E., Massey,E., Mawhinney,E., McLeod,M.P., Meador,M.,
 Mel,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
 Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N.,
 Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenwo,S.,
 Ogub,M., Okunnu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,
 Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,
 Quiles,M., Ren,Y., Rivers,M., Rojas,A., Rojubenok,I., Rolfe,M.,
 Rull,S., Savery,G., Scherer,S., Scott,G., Shen,K., Shoshbari,N.,
 Slisdon,I., Sodergren,E., Sonaite,T., Sparks,A., Stanley,H.,
 Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K.,
 Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N.,
 Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalob,D., Vinson,R.,
 Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,
 Watlington,S., Williams,G., Williamson,A., Wleczky,R., Wooden,S.,
 Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
 Weinstein,G. and Gibbs,R.

Direct Submission

Unpublished
 2 (bases 1 to 179877)

Worley,K.C.

Submitted (09-AUG-2001) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA

Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

Project Information

Center project name: HDBP

Center clone name: RP11-340E6

Summary Statistics

Assembly program: Phrap; version 0.990329

Consensus quality: 179371 bases at least Q40

Consensus quality: 185709 bases at least Q30

Consensus quality: 188640 bases at least Q20

Estimated insert size: 180121; sum-of-coverage estimation

Quality coverage: 0x in Q20 bases; agarose-gel estimation

Quality coverage: 3.7x in Q20 bases; sum-of-coverage estimation

NOTE: Estimated insert size may differ from sequence length

(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).

NOTE: This is a 'working draft' sequence. It currently

consists of 32 contigs. The true order of the pieces

is not known and their order in this sequence record is

arbitrary. Gaps between the contigs are represented as

runs of N, but the exact sizes of the gaps are unknown.

This record will be updated with the finished sequence

as soon as it is available and the accession number will

be preserved.

1 11845: contig of 11845 bp in length
 11846 11945: gap of unknown length
 11946 11945: contig of 7616 bp in length
 11946 11945: gap of unknown length
 11946 11945: contig of 8366 bp in length
 11946 11945: gap of unknown length
 11946 11945: contig of 8786 bp in length
 11946 11945: gap of unknown length
 11946 11945: contig of 10331 bp in length
 11946 11945: gap of unknown length
 11946 11945: contig of 7468 bp in length
 11946 11945: gap of unknown length
 11946 11945: contig of 6242 bp in length
 11946 11945: gap of unknown length
 11946 11945: contig of 6648 bp in length
 11946 11945: gap of unknown length
 11946 11945: contig of 7250 bp in length
 11946 11945: gap of unknown length
 11946 11945: contig of 7317 bp in length
 11946 11945: gap of unknown length
 11946 11945: contig of 7533 bp in length
 11946 11945: gap of unknown length
 11946 11945: contig of 7105 bp in length
 11946 11945: gap of unknown length
 11946 11945: contig of 6264 bp in length
 11946 11945: gap of unknown length
 11946 11945: contig of 6309 bp in length
 11946 11945: gap of unknown length
 11946 11945: contig of 5379 bp in length
 11946 11945: gap of unknown length
 11946 11945: contig of 5909 bp in length
 11946 11945: gap of unknown length
 11946 11945: contig of 5865 bp in length
 11946 11945: gap of unknown length
 11946 11945: contig of 6235 bp in length
 11946 11945: gap of unknown length
 11946 11945: contig of 4203 bp in length
 11946 11945: gap of unknown length
 11946 11945: contig of 4930 bp in length
 11946 11945: gap of unknown length
 11946 11945: contig of 3882 bp in length
 11946 11945: gap of unknown length
 11946 11945: contig of 3222 bp in length
 11946 11945: gap of unknown length
 11946 11945: contig of 3261 bp in length
 11946 11945: gap of unknown length
 11946 11945: contig of 2620 bp in length
 11946 11945: gap of unknown length
 11946 11945: contig of 3483 bp in length
 11946 11945: gap of unknown length
 11946 11945: contig of 3181 bp in length
 11946 11945: gap of unknown length
 11946 11945: contig of 3474 bp in length
 11946 11945: gap of unknown length
 11946 11945: contig of 2048 bp in length
 11946 11945: gap of unknown length
 11946 11945: contig of 3091 bp in length
 11946 11945: gap of unknown length
 11946 11945: contig of 2126 bp in length
 11946 11945: gap of unknown length
 11946 11945: contig of 2679 bp in length
 11946 11945: gap of unknown length
 11946 11945: contig of 2038 bp in length.

FEATURES

source

Location/Qualifiers

1.179877

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="3q"

/clone="RP11-340E6"

BASE COUNT

ORIGIN

48842 a 36893 c 40248 g 50788 t 3106 others

```

Query Match          60.4%; Score 455.4; DB 2; Length 179877;
Best Local Similarity 96.5%; Pred. No. 3e-117;
Matches 520; Conservative 0; Mismatches 11; Indels 8; Gaps 5;

QY 223 GTGCTGCGGGAGATGATATCATTCCTTCACAACTCTGATGTCCCTGTGGTGA 282
    |||||||
Db 65845 GTGCTGCGGGAGATGATATCATTCCTTCACAACTCTGATGTCCCTGTGGTGA 65904

QY 283 CAGAACATCGCTCCCTCTGCGAGAGACTGGGTGTCACAAAGATGGCAATGA 342
    |||||||
Db 65905 CAGAACATCGCTCCCTCTGCGAGAGACTGGGTGTCACAAAGATGGCAATGA 65964

QY 343 GAGTGTAGAGAGCTTCTTACGCCAGTTGTCAGAGCCCTGGTGTATCTGATCCAG 402
    |||||||
Db 65965 GAGTGTAGAGAGCTTCTTACGCCAGTTGTCAGAGCCCTGGTGTATCTGATCCAG 46024

QY 403 TGCCCTCAAGAGCTTACAGCTACAGGGATGTAATTCCTCAATAGAGCCGATGATC 462
    |||||||
Db 66025 TGCCCTCAAGAGCTTACAGCTACAGGGATGTAATTCCTCAATAGAGCCGATGATC 66084

QY 463 TGTGTC--TTTGACATCATCAAGCCCTGGTTAGC--ATTGTCAGTTTATCTTACA 518
    |||||  |||||
Db 66085 TGTGTCCTTTGGAGCATATCAAGCCCTGGTTAGCATTTTGTACAGTTTATCTTACA 66144

QY 519 AATTCCTGTGATTAAGAGATTAATTAATTAAGGTGCTCTCTACCTGTGTGTG 578
    |||||||
Db 66145 AATTCCTGTGATTAAGAGATTAATTAATTAAGGTGCTCTCTACCTGTGTGTG 66204

QY 579 TGTGCGGACACAGCTTGAAGTGTATTAAGAAAGAGCTCCAAATGATGATC 637
    |||||||
Db 66205 TGTGCGGACAC--CAGCTTGAAGTGTATTAAGAAAGAGCTCCAAATGATGATC 66263

QY 638 TTATATTTTACCATTTCTATACACAGGAGCTGGAAGCAGTTTC--GAGACTTTTTCGA 695
    |||||||
Db 66264 TTATATTTTACCATTTCTATACACAGGAGCTGGAAGCAGTTTCGAGAGACTTTTTCGA 66323

QY 696 TGCTTAATGTTGATCAGTTTAAAAAGATGTTTACAGTACAAATTAAGTCAGTTTAAA 754
    |||||||
Db 66324 TGCTTAATGTTGATCAGTTTAAAAAGATGTTTACAGTACAAATTAAGTCAGTTTAAA 66382

RESULT 10
AC067952/c 191221 bp DNA linear HTG 07-JUL-2000
LOCUS Homo sapiens chromosome 3 clone RP11-340E6, WORKING DRAFT SEQUENCE,
DEFINITION 34 unordered pieces.
ACCESSION AC067952
VERSION AC067952.5 GI:8569941
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
            Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 191221)
AUTHORS Waterston,R.H.
TITLE The sequence of Homo sapiens clone
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 191221)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (27-APR-2000) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT On Jun 16, 2000 this sequence version replaced gi:8099097.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
Project Information
Center project name: H_NH0340E06
----- Summary Statistics -----
Sequencing vector: M13; 100%

```

```

Sequencing vector: plasmid; 0%
Chemistry: Dye-primer ET; 100% of reads
Chemistry: Dye-terminator Big Dye; 0% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 176290 bases at least Q40
Consensus quality: 180539 bases at least Q30
Consensus quality: 182588 bases at least Q20
Insert size: 185000; agarose-fp
Insert size: 187921; sum-of-contigs
Quality coverage: 3.95 in Q20 bases; agarose-fp
Quality coverage: 3.96 in Q20 bases; sum-of-contigs

-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 34 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
-----
1 1361: contig of 1361 bp in length
1362 1461: gap of unknown length
1462 2995: contig of 1534 bp in length
2996 3096: gap of unknown length
3096 5034: contig of 1939 bp in length
5035 5135: gap of unknown length
5135 7084: contig of 1950 bp in length
7085 7185: gap of unknown length
7185 8703: contig of 1518 bp in length
8703 8802: gap of unknown length
8802 10840: contig of 2038 bp in length
10841 10940: gap of unknown length
10941 14031: contig of 3091 bp in length
14032 14131: gap of unknown length
14132 16813: contig of 2682 bp in length
16813 16913: gap of unknown length
16913 16914: gap of unknown length
16914 18942: contig of 2029 bp in length
18943 19042: gap of unknown length
19043 21820: contig of 2778 bp in length
21821 21920: gap of unknown length
21921 24942: contig of 3022 bp in length
24943 25042: gap of unknown length
25043 28487: contig of 3444 bp in length
28487 31448: gap of unknown length
31448 31548: contig of 2862 bp in length
31549 36026: gap of unknown length
36027 36128: contig of 4478 bp in length
36128 39355: gap of unknown length
39356 39455: contig of 3229 bp in length
39456 43710: gap of unknown length
43711 43810: contig of 4255 bp in length
43811 47963: gap of unknown length
47964 48063: contig of 4153 bp in length
48064 54298: gap of unknown length
54299 54398: contig of 6235 bp in length
54399 60263: gap of unknown length
60264 60363: contig of 5865 bp in length
60364 66617: gap of unknown length
66618 66717: contig of 6254 bp in length
66718 72626: gap of unknown length
72627 72726: contig of 5909 bp in length
72727 78806: gap of unknown length
78807 78906: contig of 6080 bp in length
78907 86036: gap of unknown length
86037 86136: contig of 7130 bp in length
86137 86377: gap of unknown length
86378 93777: gap of 7541 bp in length
93778 100883: contig of 7106 bp in length
100884 100983: gap of unknown length
100984 107226: contig of 6243 bp in length
107227 107326: gap of unknown length
107327 114577: contig of 7251 bp in length

```

```
* 114578 114677: gap of unknown length
* 114678 122145: contig of 7468 bp in length
* 122146 122245: gap of unknown length
* 122246 129861: contig of 7616 bp in length
* 129862 129961: gap of unknown length
* 129962 140573: contig of 10612 bp in length
* 140574 140673: gap of unknown length
* 140674 152684: contig of 12011 bp in length
* 152685 152784: gap of unknown length
* 152785 164145: contig of 11361 bp in length
* 164146 164245: gap of unknown length
* 164246 176090: contig of 11845 bp in length
* 176091 176091: gap of unknown length
* 176191 191221: contig of 15031 bp in length.

FEATURES
Source
1. 191221
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="3"
/clone="RP11-340E6"
1. 1361
/note="assembly_name:Contig22"
misc_feature
1462. 2995
/note="assembly_name:Contig23"
misc_feature
3096. 5034
/note="assembly_name:Contig24"
misc_feature
5135. 7084
/note="assembly_name:Contig25"
misc_feature
7185. 8702
/note="assembly_name:Contig26
vector_side:left"
misc_feature
8803. 10840
/note="assembly_name:Contig27"
misc_feature
10941. 14031
/note="assembly_name:Contig28"
misc_feature
14132. 16813
/note="assembly_name:Contig29"
misc_feature
16914. 18942
/note="assembly_name:Contig30"
misc_feature
19043. 21820
/note="assembly_name:Contig31"
misc_feature
21921. 24942
/note="assembly_name:Contig32"
misc_feature
25043. 28486
/note="assembly_name:Contig33"
misc_feature
28587. 31448
/note="assembly_name:Contig34"
misc_feature
31549. 36026
/note="assembly_name:Contig35"
misc_feature
36127. 39355
/note="assembly_name:Contig36"
misc_feature
39456. 43710
/note="assembly_name:Contig37"
misc_feature
43811. 47963
/note="assembly_name:Contig38"
misc_feature
48064. 54298
/note="assembly_name:Contig39"
misc_feature
54399. 60263
/note="assembly_name:Contig40"
misc_feature
60364. 66617
/note="assembly_name:Contig41"
misc_feature
66718. 72626
/note="assembly_name:Contig42"
misc_feature
72727. 78806
/note="assembly_name:Contig43"
misc_feature
78907. 86036
/note="assembly_name:Contig44"
misc_feature
86137. 93677
/note="assembly_name:Contig45"
misc_feature
93778. 100883
/note="assembly_name:Contig46"
misc_feature
100984. 107226
```

```
/note="assembly_name:Contig47"
107327. 114577
/note="assembly_name:Contig48"
114678. 122145
misc_feature
114678. 122145
/note="assembly_name:Contig49"
122246. 129861
misc_feature
122246. 129861
/note="assembly_name:Contig50"
129962. 140573
misc_feature
129962. 140573
/note="assembly_name:Contig51"
140674. 152684
misc_feature
140674. 152684
/note="assembly_name:Contig52"
152785. 164145
misc_feature
152785. 164145
/note="assembly_name:Contig53"
164246. 176090
misc_feature
164246. 176090
/note="assembly_name:Contig54"
176191. 191221
/note="assembly_name:Contig55"

BASE COUNT 52793 a 39732 c 42051 g 53340 t 3305 others
ORIGIN
Query Match 60.4%; Score 455.4; DB 2; Length 191221;
Best Local Similarity 96.5%; Pred. No. 3.1e-117;
Matches 520; Conservative 0; Mismatches 11; Indels 8; Gaps 5;

QY 223 GGGCTGGGGAATGATATCATTCCTCCACAACTGGCATGTCCTGGGTGAA 282
|||||
DB 186791 GTGGTGGGGAATGATATCATTCCTCCACAACTGGCATGTCCTGGGTGAA 186732

QY 283 CAGAACATCGCTGCCCTCTGCCAGCAGAGCTGGGTGCCAAGAAATCGCAATGA 342
|||||
DB 186731 CAGAACATCGCTGCCCTCTGCCAGCAGAGCTGGGTGCCAAGAAATCGCAATGA 186672

QY 343 GAGTGGTTGAAGGCTTCTTACGCGAGTTCAGAGCCCTGGTGAATCTTGAATCCAG 402
|||||
DB 186671 GAGTGGTTGAAGGCTTCTTACGCGAGTTCAGAGCCCTGGTGAATCTTGAATCCAG 186612

QY 403 TGCCTCAAAAGGCTGAGACACTACAGGAGATGAATCTTCAAAATGGAGCCGATGATC 462
|||||
DB 186611 TGCCTCAAAAGGCTGAGACACTACAGGAGATGAATCTTCAAAATGGAGCCGATGATC 186552

QY 463 TGTGCTC--TTTGACTACTCAAAAGCTTGGTTAGC--ATTGCTCACTTTATCTTCACA 518
|||||
DB 186551 TGTGCTCCTTGGAGCTCATCAAAAGCTTGGTTAGCATTGCTCACTTTATCTTCACA 186492

QY 519 AATCTCTGTGATTAAGAATATTTATTAAGGTGCTTCCACTGCTGTGTG 578
|||||
DB 186491 AATCTCTGTGATTAAGAATATTTATTAAGGTGCTTCCACTGCTGTGTG 186432

QY 579 TGTCCGACACAGCTTAAGAGTCTATATAAAGGAAAGAGCTCCAAATTCATCAC 637
|||||
DB 186431 TGTCCGACAC--CAGCTTAAGAGTCTATATAAAGGAAAGAGCTCCAAATTCATCAC 186373

QY 638 TTATATTTTACCATTCTTATACACAGCAGTGAAGCAGTTTC--GAGACTTTTTCGA 695
|||||
DB 186372 TTATATTTTACCATTCTTATACACAGCAGTGAAGCAGTTTCGAGACTTTTTCGA 186313

QY 696 TCGTTATGTTGATCATCAAGTAAAGAAAGTATACATCAATTAAGTCAAGTTTAA 754
|||||
DB 186312 TCGTTATGTTGATCATCAAGTAAAGAAAGTATACATCAATTAAGTCAAGTTTAA 186254

RESULT 11
AC060766/c 169675 bp DNA linear PRI 01-DEC-2001
LOCUS Homo sapiens chromosome 17, clone RP11-686D22, complete sequence.
DEFINITION AC060766
ACCESSION AC060766 GI:16931016
VERSION AC060766.8
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
```

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS

1 (bases 1 to 169675)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens chromosome 17, clone RP11-686D22
Unpublished
2 (bases 1 to 169675)

Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
Boguslavsky,L., Bouckhalter,B., Brown,A., Burkett,G.,
Campomori,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
Collamore,A., Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domino,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heatford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,D., Karatas,A.,
Klein,J., Larocque,K., Lamazares,R., Landers,T., Lehoccky,J.,
Levine,R., Lien,C., Liu,G., Locke,K., Macdonald,P., Marguis,N.,
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheters,R.,
Meldrum,J., Menus,L., Mihova,T., Miranda,C., Mienga,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rotman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Strange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tefaye,S., Theodore,J., Tirrell,A., Travers,M., Trigillo,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.

TITLE
JOURNAL
REFERENCE
AUTHORS

Submitted (20-APR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 169675)

Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Bouckhalter,B.,
Brown,A., Camarata,J., Campomori,A., Chang,J., Chararo,B.,
Choepel,Y., Colangelo,M., Collins,S., Collamore,A., Cook,A.,
Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gird,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Heatford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Kamat,A., Karatas,A., Kells,C., Larocque,K.,
Lamazares,R., Landers,T., Lehoccky,J., Levine,R., Liu,G.,
Maclean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., McPheters,R., Meldrum,J.,
Menus,L., Mihova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Roselli,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Spencer,B., Strange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tefaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigillo,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE
JOURNAL
COMMENT

Submitted (01-DEC-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smit,A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research
Genome Center
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
Project Information
Center project name: 16276
Center clone name: 686_D_22

FEATURES
source

Location/Qualifiers
1..169675
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="17"

/map="17"
/clone="RP11-686D22"
/clone_lib="RP11 Human Male BAC"
239..259
/rpt_family="AT-rich"
263..509
/rpt_family="AluUo"
complement(618..1197)
/rpt_family="LIMC4"
1198..1218
/rpt_family="(TG)n"
1758..1933
/rpt_family="AluUo"
1934..2116
/rpt_family="LIM4"
2123..2237
/rpt_family="AluUo"
2238..2262
/rpt_family="(TAAA)n"
2263..2367
/rpt_family="AluUo"
complement(2796..3097)
/rpt_family="AluUb"
3106..3161
/rpt_family="CT-rich"
3162..3338
/rpt_family="(TTC)n"
complement(3354..3672)
/rpt_family="AluSq"
4976..5489
/rpt_family="MERS1A"
5490..5664
/rpt_family="MERS1A-int"
5656..6452
/rpt_family="MERS1A-int"
6452..6747
/rpt_family="MERS7-int"
6746..7310
/rpt_family="MERS7-int"
7314..7401
/rpt_family="LTR54B"
7402..7491
/rpt_family="MERS7-int"
9428..9535
/rpt_family="LIME"
9925..10030
/rpt_family="LIME"
10325..10367
/rpt_family="LIM4c"
10819..11180
/rpt_family="LIM4c"
11178..12391
/rpt_family="LIM4c"
complement(12392..12694)
/rpt_family="AluUo"
12695..12868
/rpt_family="LIM4c"
13176..13320
/rpt_family="CTTAA)n"
complement(13647..13998)
/rpt_family="LIT1A1"
15861..16165
/rpt_family="AluSc"
complement(16442..16773)
/rpt_family="AluXs"
complement(17624..17785)
/rpt_family="MLTK"
complement(17816..18567)
/rpt_family="LIME"
18588..18815
/rpt_family="AluXs"
18822..18858
/rpt_family="(CAA)n"

Query Match 44.9%; Score 338.8; DB 9; Length 342;
 Best Local Similarity 99.4%; Pred. No. 6.2e-85;
 Matches 340; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGCCGACGTGGAAGAGAGAGAAACCTGGGCCCTGGCCCTCTACCTCGGGAGATGCA 60
 Db 1 ATGGCCGACGTGGAAGAGAGAGAAACCTGGGCCCTGGCCCTCTCTCTCTCGGGAGATGCA 60

QY 61 GGCTCCAGTGGGAGGCGACAAGATGTTCTCCCTCAAGAAGTGAACGGCGTGGCCATG 120
 Db 61 GGCTCAACGTCGGAGGCGACAAGATGTTCTCCCTCAAGAAGTGAACGGCGTGGCCATG 120

QY 121 TGGAGCTGGGACGTGGAGTGCATAGCTGGCCATCTGCAGGGTCCAGGTGATGATGCC 180
 Db 121 TGGAGCTGGGACGTGGAGTGCATAGCTGGCCATCTGCAGGGTCCAGGTGATGATGCC 180

QY 181 TGCTTAGATGTCAGCTGAAACAAACAGAGACTGTTGTGCTGGGAGAAATGT 240
 Db 181 TGCTTAGATGTCAGCTGAAACAAACAGAGACTGTTGTGCTGGGAGAAATGT 240

QY 241 AATCATTCTTCCACAACCTCTGCATGCTCCCTGGGTAAACAGAACATCGCTGCCCT 300
 Db 241 AATCATTCTTCCACAACCTCTGCATGCTCCCTGGGTAAACAGAACATCGCTGCCCT 300

QY 301 CTCTGCACGACGACGTGGGTGCTCCAAAGATGGGCAATGA 342
 Db 301 CTCTGCACGACGACGTGGGTGCTCCAAAGATGGGCAATGA 342

RESULT 13
 AX282592 342 bp DNA Linear PAT 02-NOV-2001
 LOCUS AX282592
 DEFINITION Sequence 7 from Patent WO0175145.
 ACCESSION AX282592
 VERSION AX282592.1 GI:16609675
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (sites)
 AUTHORS Issakani, S.D., Huang, J., Sheung, J. and Pray, T.R.
 TITLE Ubiquitin ligase assay
 JOURNAL Patent: WO 0175145-A7 11-OCN-2001;
 Rigel Pharmaceuticals, Inc. (us)
 FEATURES
 source 1..342
 location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"

BASE COUNT 82 a 84 c 108 g 68 t
 ORIGIN

Query Match 44.7%; Score 337.2; DB 6; Length 342;
 Best Local Similarity 99.1%; Pred. No. 1.7e-84;
 Matches 339; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGGCCGACGTGGAAGAGAGAGAAACCTGGGCCCTGGCCCTCTACCTCGGGAGATGCA 60
 Db 1 ATGGCCGACGTGGAAGAGAGAGAAACCTGGGCCCTGGCCCTCTCACTCGGGAGATGCA 60

QY 61 GGCTCCAGTGGGAGGCGACAAGATGTTCTCCCTCAAGAAGTGAACGGCGTGGCCATG 120
 Db 61 GGCTCAACGTCGGAGGCGACAAGATGTTCTCCCTCAAGAAGTGAACGGCGTGGCCATG 120

QY 121 TGGAGCTGGGACGTGGAGTGCATAGCTGGCCATCTGCAGGGTCCAGGTGATGATGCC 180
 Db 121 TGGAGCTGGGACGTGGAGTGCATAGCTGGCCATCTGCAGGGTCCAGGTGATGATGCC 180

QY 181 TGCTTAGATGTCAGCTGAAACAAACAGAGACTGTTGTGCTGGGAGAAATGT 240
 Db 181 TGCTTAGATGTCAGCTGAAACAAACAGAGACTGTTGTGCTGGGAGAAATGT 240

Db 181 TGCTTAGATGTCAGCTGAAACAAACAGAGACTGTTGTGCTGGGAGAAATGT 240

QY 241 AATCATTCTTCCACAACCTCTGCATGCTCCCTGGGTAAACAGAACATCGCTGCCCT 300
 Db 241 AATCATTCTTCCACAACCTCTGCATGCTCCCTGGGTAAACAGAACATCGCTGCCCT 300

QY 301 CTCTGCACGACGACGTGGGTGCTCCAAAGATGGGCAATGA 342
 Db 301 CTCTGCACGACGACGTGGGTGCTCCAAAGATGGGCAATGA 342

RESULT 14
 AF092877 1140 bp mRNA linear ROD 24-JUL-2001
 LOCUS AF092877
 DEFINITION Mus musculus zinc RING finger protein SAG mRNA, complete cds.
 ACCESSION AF092877
 VERSION AF092877.1 GI:4588031
 KEYWORDS
 SOURCE house mouse.
 ORGANISM Mus musculus.
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Mus.
 REFERENCE 1 (bases 1 to 1140)
 AUTHORS Swaroop, M., Bian, J., Aviram, M., Duan, H., Bisgaier, C.L., Loo, J.A. and Sun, Y.
 TITLE Expression, purification, and biochemical characterization of SAG, a RING finger redox-sensitive protein
 JOURNAL Free Radical Biol. Med. 27, 193-202 (1999)
 REFERENCE 2 (bases 1 to 1140)
 AUTHORS Sun, Y.
 TITLE Alterations of SAG mRNA in human cancer cell lines: requirement for the RING finger domain for apoptosis protection
 JOURNAL Carcinogenesis. 20 (10), 1899-1903 (1999)
 REFERENCE 3 (bases 1 to 1140)
 MEDLINE 99182502
 PUBMED 10506102
 TITLE Duan, H., Wang, Y., Aviram, M., Swaroop, M., Loo, J.A., Bian, J., Tian, Y., Mueller, T., Bisgaier, C.L. and Sun, Y.
 JOURNAL SAG, a novel zinc RING finger protein that protects cells from apoptosis induced by redox agents
 REFERENCE 4 (bases 1 to 1140)
 MEDLINE 10082581
 PUBMED 10082581
 TITLE Swaroop, M., Wang, Y., Miller, P., Duan, H., Jatke, T., Madore, S.J. and Sun, Y.
 JOURNAL Yeast homolog of human SAG/ROC2/Rbx2/Hrt2 is essential for cell growth, but not for germination: chip profiling implicates its role in cell cycle regulation
 REFERENCE 5 (bases 1 to 1140)
 MEDLINE 10851089
 PUBMED 20309864
 TITLE Duan, H., Tsvetkov, L.M., Liu, Y., Song, Y., Swaroop, M., Wen, R., Kun, H.F., Zhang, H. and Sun, Y.
 JOURNAL Promotion of S-phase entry and cell growth under serum starvation by SAG/ROC2/Rbx2/Hrt2, an E3 ubiquitin ligase component: association with inhibition of p27 accumulation
 REFERENCE 6 (bases 1 to 1140)
 MEDLINE 1125262
 PUBMED 21152847
 TITLE Molecular carcinogenesis. 30 (1), 37-46 (2001)
 JOURNAL Location/Qualifiers
 source 1..1140
 /organism="Mus musculus"
 /strain="C57BL/6xCBA"
 /db_xref="taxon:10090"
 /sex="female"
 /tissue_type="lung"


```

CDS
/dev_stage="6-8 weeks"
17..358
/Note="redox sensitive, metal binding; expression protects
cells from apoptosis induced by redox compounds"
/codon_start=1
/product="zinc RING finger protein SAG"
/protein_id="AAD25961.1"
/db_xref="GI:458032"
/translation="MADVEDGEPEVLSSHSGSAGSKSGDKMFLSKKNVAVMSMD
VECDICATCRVQVMDACLRQCAENKQEDCVVWGBCNHSHNCMSLWKNKRPCLC
QODWVVRIGK"
BASE COUNT      302 a      224 c      287 g      327 t
ORIGIN
Query Match      44.3% Score 334; DB 10; Length 1140;
Best Local Similarity 70.6%; Pred. No. 1.9e-83;
Matches 534; Conservative 0; Mismatches 180; Indels 42; Gaps 5;

QY 1 ATGGCCGACGTGGAGACGAGAGAAACCTGGCCCTGCTCTACTCCGGAGACTCA 60
    |||||
Db 17 ATGGCCGACGTGGAGACGAGAGAAACCTGGCTCTTCTTCTCGCACTCCGGAGCGCA 76
    |||||

QY 61 GGCTCCAGTCGGGAGCGCAGAGATGTTCTCCCTCAAGAAGTGAACGGGTGGCCATG 120
    |||||
Db 77 GGCTCCAGTCGGGAGCGCAGAGATGTTCTCTCAAGAAGTGAACGGGTGGCCATG 136
    |||||

QY 121 TGGAGCTGGAGCTGGAGTGGCGATACGTGGCCATCTGCAGGGTCCAGGATGATGCC 180
    |||||
Db 137 TGGAGCTGGAGCTGGAGTGGCGATACGTGGCCATCTGCAGGGTCCAGGATGATGCC 196
    |||||

QY 181 TGTCTAGATGTCAGACTGAAACAACAAGAGAGACTGTGTGTGGTGGAGATGT 240
    |||||
Db 197 TGGCTTCGATGTAACTGAAACAACAAGAGAGACTGTGTGTGGTGGAGATGT 256
    |||||

QY 241 AATCATTCCTCCACAACATGCTGATGCTGCTGGGTGAACAGAACATCGCTCCCT 300
    |||||
Db 257 AACCATTCCTCCACAACATGCTGATGCTGCTGGGTGAACAGAACATCGCTCCCT 316
    |||||

QY 301 CTCTGCACAGAGACTGGGGTCCAAAGAAATGCAAAATGAGAGTGTAGAGGCTTC 360
    |||||
Db 317 CTCTGCACAGAGACTGGGGTCCAAAGAAATGCAAAATGAGAGTGTAGAGGCTTC 376
    |||||

QY 361 TTAGCCGAGTGTTCAGAGCCCTGGTGAATCTGTATCCAGTCCCTACAAAGGCTAGA 420
    |||||
Db 377 CTGGTGTGGTGTCT-----GACCTGGCAAAAGACTAA 409
    |||||

QY 421 ACACCTACAGGGGATGTAATCTTCAATAGAGCCGATGATCTGTGCTTTGGACTCAT 480
    |||||
Db 410 ACACCTACAGGGGATGTAATCTTCAATAGAGCCGATGATCTGTGCTTTGGACTCAC 467
    |||||

QY 481 CAAGGCTGGTGT--AGCATTTGTCAGTTTATCTTCAGAAATCTCTGTGATTAAGAG 538
    |||||
Db 468 CAAGGCTGGTGTAAATTTATTTGTTAGTTTGGAAATCTCTTACAAATTT---AAG 524
    |||||

QY 539 AATAATTTAAAGTGTGCTCTCTACCTCTGTGTTGTGTGTGGGCAACACAGCTTAGA 598
    |||||
Db 525 AATAATTTAAAGTGTGCTCTCTACCTCTGTGTTGTGTGTGGGCAACACAGCTTAGA 576
    |||||

QY 599 AGTGTATATAAAGAGAGAGAGCTCAATTTGATCATCCTTAATTTATCCATTTCTAT 658
    |||||
Db 577 ATGCAATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 636
    |||||

QY 659 ACAAGAGGAGTGGAGAGCTTTTCAGACTTTTCAGTGTGTGTGTGTGTGTGTGTGT 718
    |||||
Db 637 ACATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 696
    |||||

QY 719 AAGAATGTACAGTAACAATAAAGTGCAGATTAAA 754
    |||||
Db 697 ACGAATGTACAGTAACAATAAATGCAATTGAAA 732
    |||||

RESULT 15

```

```

BC011127
LOCUS      BC011127      1125 bp      mRNA      linear      ROD 30-JUL-2001
DEFINITION Mus musculus, similar to ring finger protein 7, clone MGC:19076
IMAGE:4194107, mRNA, complete cds.
ACCESSION BC011127
VERSION   BC011127.1
KEYWORDS  GI:15029807
SOURCE    MGC.
ORGANISM  house mouse.
           Mus musculus.
           Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1125)
AUTHORS   Strausberg, R.
TITLE     Direct Submision
JOURNAL   Submitted (25-JUL-2001) National Institutes of Health, Mammalian
          Gene Collection (MGC), Cancer Genomics Office, National Cancer
          Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
          USA
REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT   Contact: MGC help desk
          Email: cga@bbs-remail.nih.gov
          Tissue Procurement: Jeffrey E. Green, M.D.
          CDNA Library Preparation: Life Technologies, Inc.
          CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
          DNA Sequencing by: Baylor College of Medicine Human Genome
          Sequencing Center
          Center code: BCM-HGSC
          Web site: http://www.hgsc.bcm.tmc.edu/cdna/
          Contact: villalob@bcm.tmc.edu.
          Villalon, D.K., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia,
          A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W.,
          Muzny, D.M., Gibbs, R.A.
          Clone distribution: MGC clone distribution information can be found
          through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov
          Series: IRAC Plate: 24 Row: k Column: 5
          This clone was selected for full length sequencing because it
          passed the following selection criteria: Genomescan gene
          prediction. Similarity but not identity to protein.
          Location/Qualifiers
            1..1125
              /organism="Mus musculus"
              /db_xref="taxon:10090"
              /clone="MGC:19076 IMAGE:4194107"
              /tissue_type="Liver, normal. 5 month old male mouse."
              /clone_lib="NCL CGAP_L19"
              /lab_host="DH10B"
              /note="Vector: pCMV-SPORT6"
            12..353
              /codon_start=1
              /product="Similar to ring finger protein 7"
              /protein_id="AAH11127.1"
              /db_xref="GI:15029808"
              /translation="MADVEDGEPEVLSSHSGSAGSKSGDKMFLSKKNVAVMSMD
              VECDICATCRVQVMDACLRQCAENKQEDCVVWGBCNHSHNCMSLWKNKRPCLC
              QODWVVRIGK"
BASE COUNT      304 a      219 c      286 g      316 t
ORIGIN
Query Match      44.2% Score 333.4; DB 10; Length 1125;
Best Local Similarity 71.3%; Pred. No. 2.8e-83;
Matches 540; Conservative 0; Mismatches 176; Indels 41; Gaps 6;

QY 1 ATGGCCGACGTGGAGACGAGAGAAACCTGGCCCTGCTCTACTCCGGAGACTCA 60
    |||||
Db 12 ATGGCCGACGTGGAGACGAGAGAAACCTGGCTCTTCTTCTCGCACTCCGGAGCGCA 71
    |||||

QY 61 GGCTCCAGTCGGGAGCGCAGAGATGTTCTCCCTCAAGAAGTGAACGGGTGGCCATG 120
    |||||
Db 72 GGCTCCAGTCGGGAGCGCAGAGATGTTCTCTCTCAAGAAGTGAACGGGTGGCCATG 131
    |||||

QY 121 TGGAGCTGGAGCTGGAGTGGCGATACGTGGCCATCTGCAGGGTCCAGGATGATGCC 180
    |||||

```

D	b		132	TGAGCTGGGACGTGTGAATGCCATTACTCTGGCATCTCGAGGGTCCAGTAGTGATGGATGCC	191
O	y		181	TGTTCTTAGATGTCAACTGAAAACAAGAAGAGACTGTGTTGTGTTGGGGAGAAATGT	240.
D	b		192	TGCCTTGATGTCAAGTGTAAAAACAAGCAAGAGAGACTGTGTTGTGTTGGGGAGAGAGT	251.
O	y		241	AATATCTCTCCACAACATCTGCATGTGCCCTGGTGGGTAAACAACAATCTGGTGCCT	300
D	b		252	AACATTTCTCTCCACAACATCTGCATGTGCCCTGGTGGGTAAACAACAATCTGGTGCCT	311.
O	y		301	CTTGCCAGCAGAGACTGGGTGTGTCCAAAGAAATCGCAATAGAGATGGTTAAGAGCTTC	360
D	b		312	CTGTGCACAGAGAGACTGGGTGTGTCCAAAGAAATCGCAATAGAGAGTGGCCAGCGCTC	371
O	y		361	TTACGCGAGTGTTCANAGCCCTGTGTGGAATCTGTAACTCACTGGCCTACAAGGCTAGA	420
D	b		372	CTGTGTGTGTG-----CTGACCCTTGACCAAGACTAA	404
O	y		421	ACACTACAGGGAGTGAATTTCTCAAAATAGAGAGCCGATGATCTGTGGTCTTGGACTCAT	480
D	b		405	ACACTGACAGGGAGTTCATCTCTGAGAGAG--AGAGATCTGTGGCCCTTGAGACTCAC	462
O	y		481	CAAAGCCTTGCT--AACATTTGTGCACTTTATCTCAGAAATTCCTGTGTATAAGAG	538
D	b		463	CAAAGCCTTGCTTTATTAATTAATTTGTCTGTTTAGTTTTGGGAAATTCCTCAATTT--AAG	519
O	y		539	ATAATTTATTAAGGTGGTCTTCTCACTCTGTGTGTGTGTGCGGCAC-ACACTTAG	597
D	b		520	ATAATTTATTAAGGTGGTGGCTTCTCACTCTGTGTGTGTGTGTGATACGATACCAATAG	579
O	y		598	AAGTCCTTAATAAAGGAGAAAGAGTCCAATTAATACACCTTTAATTTAACCACTTTCTA	657
D	b		580	AAGACCGGAACACCAAGAAATGATCTTGTTTA-----TCGTACCCACAGACGTGGG	633
O	y		658	TACACAGCGAGTGAAGSACATTTGAGACTTTTTCGATGCTTATGTGTTGATCAGTTAA	717
D	b		634	AACATTTGTTCCACAGAAACATTTGTTGTGTTTATGCTTGAGGGTTAAAAAATAGATA	693
O	y		718	AAAGATGTTACAGTAAACAATAAAGCGAGTTAA	754
D	b		694	AAGCAATGTTACAGTAAACAATAAATGACTTGA AAA 730	

Search completed: July 25, 2002, 19:14:30
Job time: 18349 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 25, 2002, 17:22:57 ; Search time 312.78 Seconds

(without alignments)
6257.697 Million cell updates/sec

Title: US-09-509-779-1

Perfect score: 1140

Sequence: 1 GTCTGCGCGCGCCCATG.....ACTAATTCATCAATTAATG 1140

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : N.Geneseq_032802.*
1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1140	100.0	1140	20	Mouse sensitive to
2	364.4	32.0	962	21	Human ORFX ORF3059
3	363.4	31.9	836	22	Human CDNA encodin
4	362.6	31.8	836	22	Human CDNA encodin
5	334	29.3	754	20	Human sensitive to
6	332.4	29.2	754	20	Human sensitive to
7	332.4	29.2	754	20	Human sensitive to
8	332.4	29.2	754	20	Human sensitive to
9	332.4	29.2	754	20	Human sensitive to

10	332.4	29.2	754	20	AAH87324	Human sensitive to
11	332.4	29.2	754	20	AAH87325	Human sensitive to
12	332.4	29.2	754	20	AAH87327	Human sensitive to
13	332.4	29.2	754	20	AAH87328	Human sensitive to
14	332.4	29.2	754	20	AAH87331	Human sensitive to
15	330.8	29.0	754	20	AAH87332	Human sensitive to
16	330.8	29.0	754	20	AAH87332	Human sensitive to
17	330.8	29.0	754	20	AAH87332	Human sensitive to
18	330.8	29.0	754	20	AAH87329	Human sensitive to
19	330.8	29.0	754	20	AAH87330	Human sensitive to
20	327.6	28.7	754	20	AAH87326	Human sensitive to
21	310	27.2	747	20	AAH87315	Human sensitive to
22	303.6	26.6	342	21	AAH88733	Nucleotide sequenc
23	303.6	26.6	342	22	AAH83570	ROC2 coding sequen
24	232.8	20.4	706	20	AAH87316	Human sensitive to
25	223.6	19.6	441	21	AAH77493	Human ORFX ORF3048
26	200.4	17.6	1152	22	AAH25847	Human apoptosis as
27	182.6	16.0	224	20	AAH41085	Human secreted pro
28	111.6	9.8	596	22	AAH25052	Human breast cance
29	108.2	9.5	169	20	AAH41482	Human secreted pro
30	102.6	9.0	630	22	AAH22339	Human breast cance
31	90.6	7.9	327	21	AAH96882	Nucleotide sequenc
32	90.6	7.9	482	21	AAH03896	Human secreted pro
33	90.6	7.9	508	21	AAH74978	DNA encoding a hum
34	90.6	7.9	5347	24	AAH94844	Human DNA sequenc
35	87.8	7.7	498	23	ABL14673	Drosophila melanog
36	87.8	7.7	10025	23	ABL14672	Drosophila melanog
37	84.8	7.4	557	22	AAH97861	Marine 7-transmemb
38	84	7.4	560	23	ABL18585	Drosophila melanog
39	84	7.4	2560	23	ABL18584	Drosophila melanog
40	84	7.4	3038	23	ABL02194	Drosophila melanog
41	84	7.4	4760	23	ABL20826	Drosophila melanog
42	84	7.4	4764	23	ABL11922	Drosophila melanog
43	83.4	7.3	3208	22	AAH12859	Human novel protei
44	81.8	7.2	692	23	ABL22527	Drosophila melanog
45	79.4	7.0	504	21	AAH74980	DNA encoding a mur

ALIGNMENTS

RESULT 1	
AAH87313	
ID	AAH87313 standard; CDNA; 1140 BP.
XX	
AC	AAH87313;
XX	
DT	27-SEP-1999 (first entry)
XX	
DE	Mouse sensitive to apoptosis (SAG) gene.
XX	
KW	SAG gene; sensitive to apoptosis; mouse; cancer; tumour;
KW	neurodegenerative disease; muscular dystrophy; wound healing;
KW	vulnerary; therapy; ds.
XX	
OS	Mus musculus.
XX	
FT	Key
FT	CDS
FT	Location/Qualifiers
FT	17..358
FT	/*tag= a
XX	
PN	MO9932514-A2.
PD	01-JUL-1999.
XX	
PF	15-DEC-1998; 98WO-US26705.
XX	
PR	11-SEP-1998; 98US-0099840.
PR	19-DEC-1997; 97US-0068179.
XX	
PA	(WARN) WARNER LAMBERT CO.
XX	
PI	Sun Y;

PD 05-OCT-2000 .
XX
PF 31-MAR-2000; 2000WO-US08621 .
XX
PR 31-MAR-1999; 99US-0127607 .
PR 02-APR-1999; 99US-0127636 .
PR 05-APR-1999; 99US-0127728 .
PR 30-MAR-2000; 2000US-0540763 .
XX
XX
PA (CURA-) CURAGEN CORP .
XX
PI Shimkets RA, Leach M;
XX
DR WPI; 2000-602362/57 .
DR P-PSDB; AAB43295 .
XX
PT Novel nucleic acids and peptides derived from open reading frame X,
PT useful for treating e.g. cancers, proliferative disorders,
XX neurodegenerative disorders and cardiovascular disease -
XX
PS Claim 5; Page 5300; 5507pp; English.

CC AAC77444 to AAC77606 encode the proteins given in AAB04027 to AAB43397.
CC which represent the human ORF open reading frames 1 to 3161. The ORF
CC sequences have activities such as: cytostatic; hepatotropic; vulnerary;
CC antiproliferative; antiparinsonian; neutropic; neuroprotective;
CC osteopathic; anticoagulant; antitarric; immunopressant;
CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
CC antidiabetic; hypotensive; dermatological; immunosuppressive;
CC antilelflammatory; antibacterial. The sequences can be used for determining
CC antitumor; and antianemic. The sequences can be used for determining
CC the presence of or predisposition to, or preventing or treating
CC pathological conditions associated with an ORF-associated disorder. The
CC nucleic acids can be used to express ORF proteins in gene therapy
CC vectors. The proteins and nucleic acids may be used to treat cancers,
CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
CC graft vs host disease, cardiovascular disease, diabetes mellitus,
CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
CC allergic, aplastic anemia, burns, wounds, bone and cartilage damage,
CC nocturnal haemoglobinuria, antilelflammatory disease; to enhance
CC coagulation; to inhibit thrombolysis; and as a contraceptive.

Query Match	32.0%;	Score 364.4;	DB 21;	Length 962;
Best Local Similarity	78.1%;	Pred. No. 3.5e-94;		
Matches 503; Conservative	0;	Mismatches 106;	Indels 35;	Gaps 4;

OY 1 GTTGGGCGCGCCCATNGGCCACGTGAGAGAGAGCGAGAAACCCTGCCTTTCTTC 60
| | | | | | | | | | | | | | | | | | | | | |
Db 36 gctccgcgycgcgcgcacatgycgagcgtggaagaagcgaaaccctgcgcctgcctc 95
| | | | | | | | | | | | | | | | | | | | | |
OY 61 GCATCCGGGAGACCAAGCTCCAAGTGGGAGGCGACAAGAATTCTCTCAAGAAGTG 120
| | | | | | | | | | | | | | | | | | | | | |
Db 96 tcaatccggagagcttcaggctccaagtctggagagggacaagaatgcttccctcaagaatg 155
| | | | | | | | | | | | | | | | | | | | | |
OY 121 GAACCGCGGTACCATGTGTGAAGCTGGACGTTGAGTGGGATTAACCTGTGCCACTTGCAGGGT 180
| | | | | | | | | | | | | | | | | | | | | |
Db 156 gaacgcggtgycgcactgtgagagcttgagagctggaatacgtgcacctgcgaaggt 215
| | | | | | | | | | | | | | | | | | | | | |
OY 181 CCAGGTGATGATCCCTCGCCTTCATGTCATCAAGCGAANAACAGAACAGACGTGTGT 240
| | | | | | | | | | | | | | | | | | | | | |
Db 216 ccaagtgatctgaaatgcccgtcctctaagaatgcaaaccaaaaaggaactgctgtcgt 275
| | | | | | | | | | | | | | | | | | | | | |
OY 241 GGTCGGGAGAGATATAACATTCTTCCACAACTGCTGATGTCTCCTGGGGTGAACA 300
| | | | | | | | | | | | | | | | | | | | | |
Db 276 ggtctggggagaaatgatatacatcttccctccacaactgctgatgctccctgtygtgtaaca 335
| | | | | | | | | | | | | | | | | | | | | |
OY 301 GAACAAATGCTGCCCTCTGTGTCCACAGAGACTGGGTAGTCCAAAAGATCGGCAATATAGA 360
| | | | | | | | | | | | | | | | | | | | | |

Db	336	gaacaatgcgtgccctctctccagcaggaactgggtgtgtccaaaagaatcgccaaltbaga	395
QY	361	GMTGGCCAGCGCGTCTCGTGTGTGTGGTCGAC-----	393
Db	396	gttgttagaaggtctcttaagcgaactgtgtccaaagaccctgtgtgattcttgatccagtg	455
QY	394	CCTGGACAAAGACTMAACACTGCAGGGGATTCATCTTGAAGAG-----AGAGATGCTG	449
Db	456	ccttacaagaagctagaacacacacagaggatgtgaattcttccaatagagcgaatgtaactg	515
QY	450	TGGCGCTTTTGAGACTGCACCAAGGCTGCTTATTATTTGTGTCGTTACTTTTGGGAAA	509
Db	516	tggtcctcttggaactcaatcaaaagccttgggttagcatttgcagtttactcttcgaaga	575
QY	510	TTTCTACAAAT---AAGATAATTTGTTAAAAATGCGCTTTCTACCTCGTGTGTGTG	566
Db	576	tctctcgtatgaagaagataattatlaaaggctgctcctccacacctgtgtgtgtg	635
QY	567	TGTGATACGAATGCATGGAAGGAGGAGCAACCGAANAATGATG	610
Db	636	tcgcgcac-acagcttagaagtgctataaaaaaagaagaagctc	678

RESULT	3
AAS26319	
ID	AAS26319 standard; cDNA; 836 BP
XX	
AC	AAS26319;
XX	
DT	07-NOV-2001 (first entry)
XX	

Human cDNA encoding a novel secreted protein, Seq ID 498.

KW Human immunosuppressive; antiarthritic; ss; antirheumatic;
KW cyclostatic; cardiant; vasotropic; cerebroprotective; nocrotropic;
KW neuroprotective; antibacterial; vincide; fungicide; ophthalmological;
KW vulnary; secreted protein; rheumatoid arthritis;
KW hyperproliferative disorder; cardiovascular disorder; cardiac arrest;
KW cerebrovascular disorder; cerebral ischaemia; anglogenesis;
KW nervous system disorder; Alzheimer's disease; infection; ocular disorder
KW corneal infection; wound healing; epithelial cell proliferation;
KW skin ageing; food additive; preservative; antiproliferative.

Homo sapiens.

W0200155322-A2

02-AUG-2001.

17-JAN-2001: 2001WO-11501341

31 - TAN-2000, 2000IF-0179065

04-FEB-2000; 2000US-0180628.

02-MAR-2000; 2000US-0186350.

16-MAR-2000; 2000US-01898/4.
17-MAR-2000: 2000US-0190076.

18-APR-2000; 2000US-0198123.
18-MAY-2000; 2000US-0305515

07-JUN-2000; 2000US-0209467.

30-JUN-2000; 2000US-0215135.

07-JUL-2000; 2000US-0216880

11 - JUL - 2000; 2000US-0217487.
11 - THU - 2000; 2000US-0217486

14-JUL-2000; 2000US-0218290.

26-JUL-2000; 2000US-0220964.

14-AUG-2000; 2000US-0224519

14 - AUG - 2000; 2000US-0225213
14 - AUG - 2000; 2000US-0225214

PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 14-AUG-2000; 2000US-0225759.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226868.
 PR 22-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232337.
 PR 14-SEP-2000; 2000US-0232338.
 PR 14-SEP-2000; 2000US-0232339.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236337.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.

PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0251988.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251989.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 (HUMA-) HUMAN GENOME SCI INC.
 Rosen CA, Barash SC, Ruben SM;
 WPI; 2001-488783/53.
 P-PSDB; AAU16332.
 DR New nucleic acid molecules encoding 461 human secreted proteins for
 PT diagnosing, preventing, treating or ameliorating medical conditions and
 PT used as food additives or preservatives -
 XX
 PS Claim 1; SEQ ID No 498; 980bp; English.
 XX
 CC The invention relates to isolated nucleic acid molecules and their
 CC encoded secreted proteins. The nucleic acids and proteins are used to
 CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
 CC rabbits, goats, horses, cats, dogs, chickens or sheep. They
 CC are also used in diagnosing a pathological condition or susceptibility
 CC to a pathological condition. Antibodies to the proteins can also
 CC be used in alleviating symptoms associated with the disorders and in
 CC diagnostic immunoassays e.g. radioimmunoassays or enzyme linked
 CC immunosorbent assays (ELISA). Disorders which are diagnosed or treated
 CC include autoimmune diseases e.g. rheumatoid arthritis,
 CC hyperproliferative disorders e.g. neoplasms of the breast or liver,
 CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders
 CC e.g. cerebral ischemia, angiogenesis, nervous system disorders e.g.
 CC Alzheimer's disease, infections caused by bacteria, viruses and fungi
 CC and ocular disorders e.g. corneal infection, and many other
 CC disorders listed in the specification. The polypeptides can also
 CC be used to aid wound healing and epithelial cell proliferation, to
 CC prevent skin aging due to sunburn, to maintain organs before

PR	25-SEP-2000	2000US -0234.998
PR	26-SEP-2000	2000US -0234.998
PR	27-SEP-2000	2000US -0235.484
PR	27-SEP-2000	2000US -0235.836
PR	27-SEP-2000	2000US -0235.836
PR	29-SEP-2000	2000US -0236.327
PR	29-SEP-2000	2000US -0236.327
PR	29-SEP-2000	2000US -0236.368
PR	29-SEP-2000	2000US -0236.370
PR	02-OCT-2000	2000US -0237.037
PR	02-OCT-2000	2000US -0237.038
PR	02-OCT-2000	2000US -0237.038
PR	02-OCT-2000	2000US -0237.040
PR	13-OCT-2000	2000US -0239.935
PR	13-OCT-2000	2000US -0239.937
PR	20-OCT-2000	2000US -0240.960
PR	20-OCT-2000	2000US -0241.221
PR	20-OCT-2000	2000US -0241.785
PR	20-OCT-2000	2000US -0241.785
PR	20-OCT-2000	2000US -0241.787
PR	20-OCT-2000	2000US -0241.808
PR	20-OCT-2000	2000US -0241.808
PR	20-OCT-2000	2000US -0241.826
PR	01-NOV-2000	2000US -0244.617
PR	08-NOV-2000	2000US -0246.474
PR	08-NOV-2000	2000US -0246.475
PR	08-NOV-2000	2000US -0246.476
PR	08-NOV-2000	2000US -0246.476
PR	08-NOV-2000	2000US -0246.477
PR	08-NOV-2000	2000US -0246.478
PR	08-NOV-2000	2000US -0246.523
PR	08-NOV-2000	2000US -0246.523
PR	08-NOV-2000	2000US -0246.525
PR	08-NOV-2000	2000US -0246.525
PR	08-NOV-2000	2000US -0246.527
PR	08-NOV-2000	2000US -0246.528
PR	08-NOV-2000	2000US -0246.532
PR	08-NOV-2000	2000US -0246.609
PR	08-NOV-2000	2000US -0246.610
PR	08-NOV-2000	2000US -0246.610
PR	08-NOV-2000	2000US -0246.611
PR	08-NOV-2000	2000US -0246.613
PR	17-NOV-2000	2000US -0249.217
PR	17-NOV-2000	2000US -0249.218
PR	17-NOV-2000	2000US -0249.218
PR	17-NOV-2000	2000US -0249.224
PR	17-NOV-2000	2000US -0249.224
PR	17-NOV-2000	2000US -0249.225
PR	17-NOV-2000	2000US -0249.226
PR	17-NOV-2000	2000US -0249.265
PR	17-NOV-2000	2000US -0249.297
PR	17-NOV-2000	2000US -0249.329
PR	17-NOV-2000	2000US -0249.330
PR	01-DEC-2000	2000US -0250.160
PR	01-DEC-2000	2000US -0250.161
PR	01-DEC-2000	2000US -0250.330
PR	05-DEC-2000	2000US -0251.988
PR	05-DEC-2000	2000US -0251.988
PR	05-DEC-2000	2000US -0256.719
PR	06-DEC-2000	2000US -0251.479
PR	06-DEC-2000	2000US -0251.856
PR	08-DEC-2000	2000US -0251.868
PR	08-DEC-2000	2000US -0251.868
PR	08-DEC-2000	2000US -0251.869
PR	08-DEC-2000	2000US -0251.989
PR	08-DEC-2000	2000US -0251.990
PR	11-DEC-2000	2000US -0251.997

Query Match	Best Local Match	Similarity	Score	31.8%	DB 22	Length 836	
Matches	500	Conservative	78.2%	0	Mismatches 104	Indels 35	Gaps
0Y	6	GGCGCCGCGCCATGGCCGAGCTGGAGGAGCGGAGGAACCTCGCTCTTCTTGCGACT	65				
Db	5	gggagcgccgcatcggcgagcgtggaagcgaggaggaagcaacgtcgccgtcctcact	64				
0Y	66	CCGGAGGCGCAGGCTTCCAAAGTGGGAGGCGACAAAGATGTTCTCTCAAGAAGTGAACG	125				
Db	65	ccggagagctcagagctcccaagtcggagagcgcaaaagatgtctccctcaagaagtggaa	124				
0Y	126	CGGTGACCATGGAGGCTGGAGCGGAGCTTGAATGGGATACCTGTCATCTCAGAGGTTCCAG	185				
Db	125	cggtygcacatgcygagcgctcggagcgtggaatgagatacgtgcgcactctcaggggtccagg	184				
0Y	186	TGATGATGCTCTGCTTGGATGTCGAAGCTGAAACAAAGCAGAAGAGACTGTGTGTGCTCT	245				
Db	185	tgatgatgctcgtctcttagatgtatcaagctgtaaaaaaacaagaagagactgtgtgtgctc	244				
0Y	246	GGGAGAGGATGATACCATTCCTCCCAACATCGTGGATGTCCTGTGGGCGGAACGAACA	305				
Db	245	ggggagagaaatgaatacattccctcccaacatgctcatgtccctctggygagaaacgaagaa	304				
0Y	306	ATTCGCTGCCCTCTGTGACGACGAGACTGGGTAGTCCAAAAGATTCGGAATGAGAGTGG	365				
Db	305	atcgtctgcctctctcgcagcagagactggtgtgtctcaagaatcggaatgagagtgtgt	364				
0Y	366	CCCAGGCGCTCCTGGTGTGTGCTGAC-----CCTGG	398				
Db	365	tagaagagctctttagacgaagttgttcaagagccctggtgtatctgtatcatcagtgcccta	424				
0Y	399	ACAAAGACTAAACATCGCAGGAGGATTCATCCTTGAGAGAG-----AGAGATCCTGTGCGC	454				

DB 425 caaaggctagaaacacacacagggagatgaatcctcaaatggagccgagatcgtgctc 484
 QY 455 CTTTGAGACTCACCAAGGCTGCTTATTATTGTTGTTGTTGGAATTCCTC 514
 DB 485 ctttggaactcaacaagccttggttagcatlctgcaattcattcagaaattc 544
 QY 515 TACAATT---AAGATAATTGTTAAAGGCTTCCCTACCTGCTGCTGTGTGA 571
 DB 545 tgcgattaagaagataattatcaagaagtgctcctcctcctcgtgtgtgtcgcg 604
 QY 572 TACGAATGCATAGAGAGCAGACACAGAAATGATC 610
 DB 605 cac-acagcttagaagtgctataaaaaaggaagagctc 642

RESULT 5
 AAX87314
 ID AAX87314 standard; cDNA: 754 BP.
 XX
 AC AAX87314;
 DT 27-SEP-1999 (first entry)
 DE Human sensitive to apoptosis (SAG) gene.
 XX
 KW SAG gene; sensitive to apoptosis; human; cancer; tumour;
 neurodegenerative disease; muscular dystrophy; wound healing;
 XX
 KW vulnerability; therapy; ds.
 XX
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT CDS 1..342
 FT /ftag a
 PN W0932514-A2.
 PD 01-JUL-1999.
 XX
 PF 15-DEC-1998; 98WO-US26705-
 PR 11-SEP-1998; 98US-0099840.
 PR 19-DEC-1997; 97US-0068179.
 XX
 PA (WARN) WARNER LAMBERT CO.
 PI Sun Y;
 DR WPI: 1999-430152/36.
 DR P-PSDB; AAY06492.
 PT SAG: Sensitive to Apoptosis Gene and related proteins, useful for
 PT promoting cell growth and protecting cells against apoptosis
 XX
 PS Claim 9; Page 50-51; 84pp; English.
 XX
 CC This is the nucleotide sequence of human sensitive to apoptosis
 CC (SAG) cDNA, which codes for a novel redox-sensitive, haem-binding
 CC protein (see AAY06492) with a zinc RING finger domain that promotes
 CC cell growth, protects cells from apoptosis, scavenges oxygen
 CC radicals and can be used for the reversion of a tumour phenotype.
 CC The cDNA was isolated from a HeLa cell cDNA library using mouse
 CC SAG cDNA (see AAX87313) as probe. SAG is highly conserved among
 CC species. Disruption in yeast was shown to be lethal. SAG deletion
 CC mutants (see AAX87315-16) have been identified in human cancer lines,
 CC suggesting a role in carcinogenesis. SAG genes, and mutant SAG
 CC genes, can be used to protect cells from apoptosis induced by redox
 CC reagents. Antisense SAG genes can be used to inhibit the growth of
 CC tumour cells. The SAG genes can also be used for the recombinant
 CC production of the SAG proteins. The SAG proteins can be used to
 CC scavenge oxygen radicals in organisms and to promote wound healing.
 CC Additionally, the SAG genes or their complements can be used to

CC promote or inhibit the growth of plant cells (all claimed). The SAG
 CC protein is also an ideal molecular target in the development of
 CC drugs against neurodegenerative disorders, cancers and muscle
 CC dystrophy.
 XX
 SQ Sequence 754 BP; 205 A; 155 C; 201 G; 193 T; 0 other;

Query Match 29.3%; Score 334; DB 20; Length 754;
 Best Local Similarity 70.6%; Pred. No. 1,6e-85;
 Matches 534; Conservative 0; Mismatches 180; Indels 42; Gaps 5;

QY 17 ATGGCCGAGCTGAGAGAGCGGCGAGAAACCTGCTCTTCTCGGACACCGGAGCGCA 75
 DB 1 atggccgagctggaagacggaaggaaccgcgcctgcctcctcaccgcggaagctca 60
 QY 77 GGCTCCAGTCGGAGAGCGGCGAGAAATGCTCTCTCAGCAAGTGAACGCGGTAGCCATG 136
 DB 61 ggtcccaagtcggagagcgagcaagatgtctcctcctcaagaagatggaagcggctgc 120
 QY 137 TGGAGCTGGAGCTTGAGTCCGATACCTGTCGATCGAGGTCAGAGTGAATGCC 196
 DB 121 tggagctggagcgtggagtgcgatcagctgcgcacatcgcagggccagtgatgagtc 180
 QY 197 TGCCTTCGATGTCAGCTGGAACAAGCAAGAGAGAGAGTGTGTGTGTGTGGAGAGTGT 236
 DB 181 tgccttagatgltcaagctgaaacaaagagagactgltgtgtgtgtgtgtgtgtgtgt 240
 QY 257 AACCAATTCCTTCACAACTGCTCATGCTCCCTGTGGTGAACAAGAACATTCCTGCC 316
 DB 241 aatcattcctccacaactctgcatgctcctgtgtgtgtgtgtgtgtgtgtgtgtgtgt 300
 QY 317 CTGTGCAGAGAGACAGTGGGATGTCGAAGAATGGCAAAAGAGAGTGGCCAGGCGCTC 376
 DB 301 ctgtgcagagagactggt 360
 QY 377 CTGCTGTGCTTG-----CTGACCTTGAGCAAAAGACTAA 409
 DB 361 ttgagcgagctgtgtcagagccctggtgactctgtatccagtcagtcctcaaaagctc 420
 QY 410 ACACTGCGAGGGATTCATCCTTGAGAGAG--AGAGGATGCTGTGGCCTTGAAGTCA 467
 DB 421 acaactcagggatgattctcctcaaaagagagagctgtgtgtgtgtgtgtgtgtgtgt 480
 QY 468 CAAGGCTTCTTATTAAATTTGCTGTAGTTTGGGAAATTCCTCAATT---AAG 524
 DB 481 caaagccttggttag--cattgtcagtttattctcctcaaatctctgtgtatgaag 538
 QY 525 ATAAATTTGTTAAAGATGCTCTTCTACCTGCTGTGTGTGT-----GTGATACGA 576
 DB 539 ataatttataaagtggtgtcctcctcctcctcctcctgtgtgtgtgtgtgtgtgtgtgt 598
 QY 577 ATGCATGAGAGAGCGAGACACAGAAATGATCTTTGTTATCTGTACCCAGCACTGA 636
 DB 599 agtgcataaaaaaaggaagagctcccaattgataccttaattaccatttctat 658
 QY 637 ACATTTGTTACAGAGAAGAACTTTGTTGTTATCTGAGGTTAAATAATAGTAA 696
 DB 659 acacagagcgtggagcagcttcagacttllcagactgtatgtgtgtgtgtgtgtgt 718
 QY 697 ACGAATGTACACTAACAATAAATGCAATTGAA 732
 DB 719 aagaagtgtcagtaacaataaagtcaggtttaa 754

RESULT 6
 AAX87317
 ID AAX87317 standard; cDNA: 754 BP.
 XX
 AC AAX87317;
 DT 27-SEP-1999 (first entry)
 XX


```

QY 137 TGGAGCTGGAGCGTTGAGTGCATACCTGTCATCTGCAGGGTCCAGTGTATGATGCC 196
    |||||
Db 121 tggagctggagcggtgagtgatgacgtgacgtccatctcgcaggggtccagtgatgagtc 180
QY 197 TGCCTTGATGTCAAGCTGAAAAACAGCAAGAGAGCTGTCTGTGTGGGAGAGTGT 256
    |||||
Db 181 tgccttgatgtcaagctgaaaaaacaagaagagctgtgtgtgtgtgtgtgtgtgtgtgtgt 240
QY 257 AACCATTCCTCCACACATCTGCATGTCCTGTGGGTGAACGACATCGTGGCCCT 316
    |||||
Db 241 aatcatctccacacactgcacatgctcctgtggtgtgaaacaacatgcctgcctc 300
QY 317 CTGTGCACGAGAGCTGGTGTAGTCCAAAGAATCGCAAAATGAGAGTGGCCAGCCCTC 376
    |||||
Db 301 ctgtgcacgagagctgggtgtgtccaaagaatcgcaaaagagagtggttagaagcttc 360
QY 377 CTGGTGTGTTG-----CTGACCTCGACCAAGACTAA 409
    |||||
Db 361 ttgagcgagtggttcagagccctgtgtgatctgttaaccagtgccctacaaggtcaga 420
QY 410 ACACCTGAGGGAGTCACTCCTTGAGAGAG--AGAGATGCTGTGCGCCTTGAGACTCAC 467
    |||||
Db 421 acactacagggagatgaaatcttcaaatagagccgagctgtgtgtgtgtgtgtgtgtgt 480
QY 468 CAAAGCGTCTTATTAATTTGCTGTGTTAGTTTGGAAATTCCTACAAAT---AAG 524
    |||||
Db 481 caaagccttgtag--cattgttcagtttacttcaagaatctctgtgtgtgtgtgtgtgt 538
QY 525 ATAAATTTGTTAAATGCGCTTCTCTACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 576
    |||||
Db 539 ataattattaaagtggtctctctacccctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 598
QY 577 ATGATGGAAGCGAGAACACCAAGAAATGATCTTTGTTATCTGATCCACGACTGGA 636
    |||||
Db 599 agtgcataaaaggaagagatcccaatgaaatgaaatccctaatcttcaatcttcaat 658
QY 637 ACATTTGTTTCACAGAGACATGTTGTTATGCTTGAGGGTTAAATAATAGATAA 696
    |||||
Db 659 acaacagagcgatggaagcagatcttcgagaacttttcgagctgtatgtgtgtgtgtgt 718
QY 697 ACGAATGTACAGTAAACAAATTAATGCAATTGAAAA 732
    |||||
Db 719 aagatgtttacagtaacaataaagtgtagtttaa 754

RESULT 10
AA87324
ID AAX87324 standard; cDNA; 754 BP.
AC AAX87324;
XX
DT 27-SEP-1999 (first entry)
XX
DE Human sensitive to apoptosis (SAG) gene mutant MM8.
XX
KM SAG gene; sensitive to apoptosis; human; cancer; tumour;
KM neurodegenerative disease; muscular dystrophy; wound healing;
KM vulnerability; therapy; mutant; ds.
XX
OS Homo sapiens.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT CDS 1..342
FT FT /*tag= a
FT FT replace(262,T)
FT FT /*tag= b
FT FT /note= "C88s mutation"
XX
XX WC09932514-A2.
XX
PD 01-JUL-1999.

```

```

XX
PF 15-DEC-1998; 98MO-US26705.
XX
PR 11-SEP-1998; 98US-0099840.
PR 19-DEC-1997; 97US-0068179.
XX
PA (WARN ) WARNER LAMBERT CO.
XX
PI Sun Y;
XX
DR WPI; 1999-430152/36.
DR P-PSDB; AAY06502.
XX
PT SAG: Sensitive to Apoptosis Gene and related proteins, useful for
XX promoting cell growth and protecting cells against apoptosis
XX
PS Claim 15; Page 71; 84pp; English.
XX
XX This is the nucleotide sequence of human sensitive to apoptosis
XX (SAG) mutant gene MM8, which codes for a SAG protein (see AAY06502)
XX in which the Cys residue at position 88 of the native protein (see
XX AAY06492) is replaced by a Ser residue owing to a mutation of codon
XX 88 from TGC to AGC obtained by site-directed mutagenesis of SAG
XX cDNA. This residue is in zinc-finger finger 2 of SAG. Single and
XX double SAG mutants (see AAX87317-31) were made in order to determine
XX the role of each cysteine residue of SAG in haem binding and SAG
XX oligomerization. These properties were unaffected by the MM8
XX mutation. SAG is a novel zinc finger protein that promotes
XX cell growth, protects cells from apoptosis, scavenges oxygen
XX radicals and can be used for the reversal of a tumour phenotype.
XX SAG genes, and mutant SAG genes, can be used to protect cells from
XX apoptosis induced by redox reagents. They can also be used for the
XX recombinant production of SAG proteins, which are molecular targets
XX in the development of drugs against neurodegenerative disorders,
XX cancers and muscle dystrophy, and promoting wound healing.
XX
XX Sequence 754 BP: 206 A; 155 C; 201 G; 192 T; 0 other;
XX
Query Match 29.2%; Score 332.4; DB 20; Length 754;
Best Local Similarity 70.5%; Pred. No. 4,7e-85;
Matches 533; Conservative 0; Mismatches 181; Indels 42; Gaps 5;
XX
QY 17 ATGGCCGACGAGGAGCGGAGAGACCTGCGCTCTTCTTCGACATCCGGAGCGCA 76
    |||||
Db 1 atggcgagctggaagcggagagaaacctgcgctctcctccctcgcggagctca 60
QY 77 GGCCTCAAGTCGGAGGCGCAGAGATGTTCTCTCAAGAAGTGAAGCGCGTAGCCATG 136
    |||||
Db 61 ggcctcaagtcggagggcgagaaagatgtctcctccaagaagtgaagcgtgtgcatg 120
QY 137 TGGAGCTGGAGCGTTGAGTGCATACCTGTGCTGTCAGAGGTCCAGGTGATGATGCC 196
    |||||
Db 121 tggagctggagcggtgagtgatgacgtgacgtccatctcgcaggggtccagtgatgagtc 180
QY 197 TGCCTTGATGTCAAGCTGAAAAACAGCAAGAGAGCTGTCTGTGTGGGAGAGTGT 256
    |||||
Db 181 tgccttgatgtcaagctgaaaaaacaagaagagctgtgtgtgtgtgtgtgtgtgtgtgtgt 240
QY 257 AACCATTCCTCCACACATCTGCATGTCCTGTGGGTGAACGACATCGTGGCCCT 316
    |||||
Db 241 aatcatctccacacactgcacatgctcctgtggtgtgaaacaacatgcctgcctc 300
QY 317 CTGTGCACGAGAGCTGGTGTAGTCCAAAGAATCGCAAAATGAGAGTGGCCAGCCCTC 376
    |||||
Db 301 ctgtgcacgagagctgggtgtgtccaaagaatcgcaaaagagagtggttagaagcttc 360
QY 377 CTGGTGTGTTG-----CTGACCTCGACCAAGACTAA 409
    |||||
Db 361 ttgagcgagtggttcagagccctgtgtgatctgttaaccagtgccctacaaggtcaga 420
QY 410 ACACCTGAGGGAGTCACTCCTTGAGAGAG--AGAGATGCTGTGCGCCTTGAGACTCAC 467
    |||||

```


XX W09932514-A2.
 PN 01-JUL-1999.
 XX 15-DEC-1998; 98WO-US26705.
 XX 11-SEP-1998; 98US-0099840.
 PR 19-DEC-1997; 97US-0068179.
 XX (WARN) WARNER LAMBERT CO.
 PA Sun Y;
 PI WPI: 1999-430152/36.
 DR P-PSDB: AAY06506.
 XX
 PT SAG: Sensitive to Apoptosis Gene and related proteins, useful for
 PT promoting cell growth and protecting cells against apoptosis
 PS Claim 15; Page 77-78; 84pp; English.

XX This is the nucleotide sequence of human sensitive to apoptosis
 CC (SAG) mutant gene MM12, which codes for a SAG protein (see AAY06506),
 CC in which the Cys residue at position 73 of the native protein (see
 CC AAY06492) is replaced by a Ser residue owing to a mutation of codon
 CC 73 from TGT to AGC obtained by site-directed mutagenesis of SAG
 CC cDNA. This residue is in a protease inhibitor motif of SAG. Single
 CC and double SAG mutants (see AX87317-31) were made to determine
 CC the role of each cysteine residue of SAG in haem binding and SAG
 CC oligomerization. These properties were unaffected by the MM12
 CC mutation. SAG is a novel zinc finger protein that promotes
 CC cell growth, protects cells from apoptosis, scavenges oxygen
 CC radicals and can be used for the reversion of a tumour phenotype.
 CC SAG genes, and mutant SAG genes, can be used to protect cells from
 CC apoptosis induced by redox reagents. They can also be used for the
 CC recombinant production of SAG proteins, which are molecular targets
 CC in the development of drugs against neurodegenerative disorders,
 CC cancers and muscle dystrophy, and promoting wound healing.

SO Sequence 754 BP; 206 A; 155 C; 201 G; 192 T; 0 other;

Query Match 29.2%; Score 332.4; DB 20; Length 754;
 Best Local Similarity 70.5%; Pred. No. 4,7e-85;
 Matches 533; Conservative 0; Mismatches 181; Indels 42; Gaps 5;
 QY 17 ATGGCCGAGCTGGAGAGCGGCGAGAACCCCTGCTCTTCTTCCGCACTCCGGAGCGCA 76
 DB 1 atggccgagctggagagcgggagagaaacctgagccttcgctcgcgagagctca 60
 QY 77 GGCCTCAAGTCGGAGGCGGCAAGATGTTCTCTCAAGAGTGGGAACCGGTAGCCATG 136
 DB 61 ggcctcaagtcggagggcgagcaagatgtcttcctcaagaagcgagctgagcctca 120
 QY 137 TGAGAGTGGGAGCTGGTGCATACCTGTCATCTGACAGGCTCAGGTGATGATGCC 196
 DB 121 tgaagctggagagctggagagctgagctgagccttcgaggtccaggtgagagctgc 180
 QY 197 TGCCCTCGAGTGAAGCTGAAGCAAGCAAGAGACTGTGTTGGTCTGGGAGAGCTG 256
 DB 181 tgccttcgagtgagctgaagctgaagcaagcaagagagagagctgtgtgtctgggagag 240
 QY 257 AACCATTCCTTCACAACTCTGCTGATGCTGCTGGTGGTAAACAGAACATGCTGCTCC 316
 DB 241 aacatcttccttcacaaactctgctgattgctgctggagagagagagagagagagag 300
 QY 317 CTGAGCCAGAGAGCTGGTAGTCAAGAAATGCGCAATGAGAGGTGGCCAGGCGCTC 376
 DB 301 ctgagccagagagagctggtagtcaagaatgcgcaatgagagaggtggtagagagcttc 360
 QY 377 CTGCTGTGTGTTG-----CTGACCTCGGCAAGACTAA 409

DB 361 tttagcgcaagtgttcaagagccctgtgtgatacttgaatcccaagtcgccaacaagcgtaga 420
 QY 410 ACATGCGAGGGATTCATCTTGAGAGAG--AGAGATGCTGTGGCCCTTGAGACTCAC 467
 DB 421 acatcagagggatgaattcttcacaaatagagagccatgagatctgtgtgtgtgagact 480
 QY 468 CAAAGGCTGCTTTATTAATTTGTTGTTAGTTTGGGAAATTCCTACAAATT--AAG 524
 DB 481 caaagcctgtgtag--cattgtgcagtttattcttcagaattctctgtgttaagaag 538
 QY 525 ATAAATTTGTTAAATGAGCCCTTCTACTCTGGTGTGTGT-----GTGATACGA 576
 DB 539 ataatatttaaggtgtgtccttcactctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 598
 QY 577 ATGATAGAAAGACGAGAACACACAGAAATGATCTTTGTTTATCTGTACCCAGACTGGA 636
 DB 599 agtgcctataaagaagaagagctccaaatgaatccattataatccattctat 658
 QY 637 ACATTTGTTTCACAGAAACATTTGTTGTTTATGCTTGAGGCTTAAATAATAGATA 696
 DB 659 acaacagcgagtgagagcagcttcgagacttttcgagtcttatgtgtgtgtgtgtgt 718
 QY 697 ACGAATGTTACGTACAAATTAATGATGATAAA 732
 DB 719 aagatgttaccagtaacaataaagtcagtttaa 754

RESULT 14
 ID AAX87331 standard; cDNA; 754 BP.
 AC AAX87331;
 XX 27-SEP-1999 (first entry)
 DT Human sensitive to apoptosis (SAG) gene mutant MM15.
 DE SAG gene; sensitive to apoptosis; human; cancer; tumour;
 KW neurodegenerative disease; muscular dystrophy; wound healing;
 KW vulnereary; therapy; mutant; ds.
 XX Homo sapiens.
 OS Synthetic.
 XX Key Location/Qualifiers
 FH 1..342
 FT /*tag= a
 FT mutation replace(139,C)
 FT /*tag= b
 FT /*note= "C47S mutation"
 W09932514-A2.
 PN 01-JUL-1999.
 XX 15-DEC-1998; 98WO-US26705.
 XX 11-SEP-1998; 98US-0099840.
 PR 19-DEC-1997; 97US-0068179.
 XX (WARN) WARNER LAMBERT CO.
 PA Sun Y;
 PI WPI: 1999-430152/36.
 DR P-PSDB: AAY06509.
 PT SAG: Sensitive to Apoptosis Gene and related proteins, useful for
 PT promoting cell growth and protecting cells against apoptosis
 PS Claim 15; Page 82-83; 84pp; English.
 XX This is the nucleotide sequence of human sensitive to apoptosis

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 25, 2002, 15:11:49 ; Search time 102.69 Seconds

(without alignments)
2726.869 Million cell updates/sec

Title: US-09-509-779-1

Perfect score: 1140

Sequence: 1 GTCTGCGCGCGCGCCATG.....ACATATCATCATTAATG 1140

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 38353 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued_Patents_NA :
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PTUS.COMB.seq:*
6: /cgn2_6/ptodata/2/ina/Dackfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40	3.5	549	1	US-07-991-867B-28
2	40	3.5	549	1	US-08-107-755A-28
3	40	3.5	549	2	US-08-544-332-28
4	40	3.5	1511	1	US-07-991-867B-8
5	40	3.5	1511	1	US-08-107-755A-8
6	40	3.5	1511	2	US-08-544-332-8
7	40	3.5	4810	3	US-08-852-629-11
8	40	3.5	4838	3	US-08-852-629-15
9	38.2	3.4	7218	1	US-08-232-463-14
10	36.4	3.2	1037	2	US-08-824-405-1
11	36	3.2	2861	1	US-08-299-953-1
12	36	3.2	2861	1	US-08-459-415-1
13	36	3.2	2861	4	US-09-066-687-1
14	36	3.2	2861	5	PCT-US95-11231-1
15	36	3.2	3881	1	US-08-299-953-2
16	36	3.2	3881	1	US-08-459-415-2
17	36	3.2	3881	4	US-09-066-687-2
18	36	3.2	3881	5	PCT-US95-11231-2
19	35.8	3.1	3019	4	US-09-359-161-2
20	35.6	3.1	3701	4	US-08-845-258-10
21	35.6	3.1	3701	4	US-08-990-571-10
22	35.6	3.1	3701	4	US-08-723-142A-10
23	35.4	3.1	4818	3	US-08-817-926-27
24	35	3.1	3181	1	US-08-655-086-1
25	34.8	3.1	767	4	US-08-998-416-472
26	34.6	3.0	591	2	US-08-943-208-1
27	34.6	3.0	1982	3	US-08-747-221B-13

28	34.6	3.0	1982	3	US-08-747-221B-15	Sequence 15, Appl
29	34.6	3.0	1982	4	US-09-005-051-13	Sequence 13, Appl
30	34.6	3.0	1982	4	US-09-005-051-15	Sequence 15, Appl
31	34.6	3.0	2144	3	US-08-747-221B-57	Sequence 57, Appl
32	34.6	3.0	2144	3	US-08-747-221B-59	Sequence 59, Appl
33	34.6	3.0	2144	4	US-09-005-051-57	Sequence 57, Appl
34	34.6	3.0	2144	4	US-09-005-051-59	Sequence 59, Appl
35	34.4	3.0	285	1	US-08-435-040-1	Sequence 1, Appl
36	34.4	3.0	285	4	US-09-020-216-1	Sequence 1, Appl
37	34.4	3.0	7904	1	US-08-316-238B-1	Sequence 1, Appl
38	34.4	3.0	7904	1	US-08-316-238B-2	Sequence 2, Appl
39	34.4	3.0	7904	1	US-08-410-005-1	Sequence 1, Appl
40	34.4	3.0	7904	3	US-08-929-140-1	Sequence 1, Appl
41	34.4	3.0	7904	4	US-09-560-579A-1	Sequence 1, Appl
42	34.2	3.0	2555	2	US-08-693-457-3	Sequence 3, Appl
43	34.2	3.0	2555	4	US-09-265-731-3	Sequence 3, Appl
44	34	3.0	3435	4	US-09-329-685A-1	Sequence 1, Appl
45	33.8	3.0	1361	4	US-09-232-191-20	Sequence 20, Appl

ALIGNMENTS

RESULT 1
US-07-991-867B-28
Sequence 28, Application US/07991867B
Patent No. 5476781
GENERAL INFORMATION:
APPLICANT: Moyer, Richard W.
APPLICANT: Hall, Richard L.
APPLICANT: Gruidl, Michael E.
TITLE OF INVENTION: No. 5476781el Entomopoxvirus Expression System
NUMBER OF SEQUENCES: 66
CORRESPONDENCE ADDRESS:
ADDRESSEE: David R. Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/991,867B
CLASSIFICATION: 435
FILING DATE: 12-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO 92/14818
FILING DATE: 12-FEB-1992
APPLICATION NUMBER: US 07/827,685
FILING DATE: 30-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/657,584
FILING DATE: 19-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanchik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: UFI14.C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 549 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
US-07-991-867B-28

Query Match

3.5%; Score 40; DB 1; Length 549;

Best Local Similarity 52.4%; Pred. No. 0.039; Mismatches 80; Indels 0; Gaps 0;

Matches 88; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

Db 941 ACCAATAGTCTTAAGAACTGTTTCTGTTTGGCGAAGTGTATGATGTTTAG 1000

Db 220 AACATACCTTTTATATATATAGCATTTTATCACAATAATGTTCTAATCATTTTC 279

Db 1001 TCAAAATATTAGTAAAGAAATGCTTACTAGTAACTGAAGTTCATTATGCAATGT 1060

Db 280 TTCAAAAATGACACATCATCTATGCCAATATATCATATATATCATGATATGATTC 339

Db 1061 TTTAATAAATATGCTTGTAGTTTAAAGTTGATATATATCTCT 1108

Db 340 ATTAATTAATATATTTGTTTAAATGATTAATATCTTTTATTAATAT 387

RESULT 2

US-08-107-755A-28

Sequence 28, Application US/08107755A

Patent No. 5721352

GENERAL INFORMATION:

APPLICANT: Moyer, Richard W.

APPLICANT: Hall, Richard L.

APPLICANT: Gruidl, Michael E.

TITLE OF INVENTION: No. 5721352el Entomopoxvirus Expression System

NUMBER OF SEQUENCES: 40

CORRESPONDENCE ADDRESS:

ADDRESSEE: David R. Saliwanchik

STREET: 2421 N.W. 41st Street, Suite A-1

CITY: Gainesville

STATE: Florida

COUNTRY: U.S.A.

ZIP: 32606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/107,755A

FILING DATE: 19-AUG-1993

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/827,658

FILING DATE: 30-JAN-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/657,584

FILING DATE: 19-FEB-1991

ATTORNEY/AGENT INFORMATION:

NAME: Saliwanchik, David R.

REGISTRATION NUMBER: 31,794

REFERENCE/DOCKET NUMBER: UF114.C2

TELECOMMUNICATION INFORMATION:

TELEPHONE: (904) 375-8100

TELEFAX: (904) 372-5800

INFORMATION FOR SEQ ID NO: 28:

SEQUENCE CHARACTERISTICS:

LENGTH: 549 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: unknown

MOLECULE TYPE: DNA (genomic)

US-08-107-755A-28

Query Match

3.5%; Score 40; DB 1; Length 549;

Best Local Similarity 52.4%; Pred. No. 0.039; Mismatches 80; Indels 0; Gaps 0;

Matches 88; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

Db 941 ACCAATAGTCTTAAGAACTGTTTCTGTTTGGCGAAGTGTATGATGTTTAG 1000

Db 220 AACATACCTTTTATATATATAGCATTTTATCACAATAATGTTCTAATCATTTTC 279

Db 1001 TCAAAATATTAGTAAAGAAATGCTTACTAGTAACTGAAGTTCATTATGCAATGT 1060

Db 280 TTCAAAAATGACACATCATCTATGCCAATATATCATATATATCATGATATGATTC 339

Db 1061 TTTAATAAATATGCTTGTAGTTTAAAGTTGATATATATCTCT 1108

Db 340 ATTAATTAATATATTTGTTTAAATGATTAATATCTTTTATTAATAT 387

RESULT 3

US-08-544-332-28

Sequence 28, Application US/08544332

Patent No. 5935777

GENERAL INFORMATION:

APPLICANT: Moyer, Richard W.

APPLICANT: Hall, Richard L.

APPLICANT: Gruidl, Michael E.

TITLE OF INVENTION: No. 5935777el Entomopoxvirus Expression System

NUMBER OF SEQUENCES: 77

CORRESPONDENCE ADDRESS:

ADDRESSEE: Gerard H. Bencen

STREET: 2421 N.W. 41st Street, Suite A-1

CITY: Gainesville

STATE: FL

COUNTRY: USA

ZIP: 32606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/544,332

FILING DATE: 07-DEC-1992

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/991,867

FILING DATE: 07-DEC-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/107,755

FILING DATE: 19-AUG-1993

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: WO 92/14818

FILING DATE: 12-FEB-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/827,685

FILING DATE: 30-JAN-1992

ATTORNEY/AGENT INFORMATION:

NAME: Bencen, Gerard H.

REGISTRATION NUMBER: 35,746

REFERENCE/DOCKET NUMBER: UF114.C4

TELECOMMUNICATION INFORMATION:

TELEPHONE: 904-375-8100

TELEFAX: 904-372-5800

INFORMATION FOR SEQ ID NO: 28:

SEQUENCE CHARACTERISTICS:

LENGTH: 549 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: unknown

MOLECULE TYPE: DNA (genomic)

US-08-544-332-28

Query Match

3.5%; Score 40; DB 2; Length 549;

Best Local Similarity 52.4%; Pred. No. 0.039; Mismatches 80; Indels 0; Gaps 0;

Matches 88; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

```

1      LOCATION: 852..1511
2      US-07-991-867B-8
3
4      Query Match      3.5%   Score 40; DB 1; Length 1511;
5      Best Local Similarity 52.4%; Pred. No. 0.065;
6      Matches 88; Conservative 0; Mismatches 80; Indels 0; Gaps 0;
7
8      Oy 941 ACCACTGAGTCTTAAGAAGCTGTTTCTGTTTGGCCGAAGTGTGATGTGTTTGG 1000
9          | | | | | | | | | | | | | | | | | | | | | | | | | | | |
10         453 AACCATACACTTTTATTTATTTATTTAGCCATTTTATTCACAAAATGTGTCMAATCATTTTC 512
11
12         Oy 1001 TCAAAATATTTAGTGGAAATGGCTTACTAGTATACACGCAAGTCTATTCGAATGT 1060
13             | | | | | | | | | | | | | | | | | | | | | | | | | | | |
14             513 TTCAAAAATATGACACTCATCTATTCACATATATATCATATATATTCAGCATATTTGATTTC 572
15
16         Oy 1061 TTTAATTAATATTTGCTGCTTGTGAGTTATTAAGTTGATATATPACTCT 1108
17             | | | | | | | | | | | | | | | | | | | | | | | | | | | |
18             Db 573 ATTAATTAATATTTGTTTATTAATGATATAATATCTTATTTAATAT 620
19
20 RESULT 5
21 US-08-107-755A-8
22 Sequence 8, Application US/08107755A
23 Patent No. 5721352
24
25 GENERAL INFORMATION:
26 APPLICANT: Moyer, Richard W.
27 APPLICANT: Hall, Richard L.
28 APPLICANT: Gruidl, Michael E.
29 TITLE OF INVENTION: No. 5721352e1 Entomopoxvirus Expression System
30 NUMBER OF SEQUENCES: 40
31 CORRESPONDENCE ADDRESSES:
32 ADDRESSEE: David R. Saliwanchik
33 STREET: 2421 N.W. 41st Street, Suite A-1
34 CITY: Gainesville
35 STATE: Florida
36 COUNTRY: U.S.A.
37 ZIP: 32606
38
39 COMPUTER READABLE FORM:
40 MEDIUM TYPE: Floppy disk
41 COMPUTER: IBM PC compatible
42 OPERATING SYSTEM: PC-DOS/MS-DOS
43 SOFTWARE: Patentin Release #1.0, Version #1.25
44 CURRENT APPLICATION DATA:
45 APPLICATION NUMBER: US/08/107,755A
46 FILING DATE: 19-AUG-1993
47 CLASSIFICATION: 435
48
49 PRIOR APPLICATION DATA:
50 APPLICATION NUMBER: US 07/827,658
51 FILING DATE: 30-JAN-1992
52 PRIOR APPLICATION DATA:
53 APPLICATION NUMBER: US 07/657,584
54 FILING DATE: 19-FEB-1991
55 ATTORNEY/AGENT INFORMATION:
56 NAME: Saliwanchik, David R.
57 REGISTRATION NUMBER: 31,794
58 REFERENCE/DOCKET NUMBER: UFI14.C2
59 TELECOMMUNICATION INFORMATION:
60 TELEPHONE: (904) 375-8100
61 TELEFAX: (904) 372-5800
62 INFORMATION FOR SEQ ID NO: 8:
63 SEQUENCE CHARACTERISTICS:
64 LENGTH: 1511 base pairs
65 TYPE: nucleic acid
66 STRANDEDNESS: double
67 TOPOLOGY: unknown
68 MOLECULE TYPE: DNA (genomic)
69 ORIGINAL SOURCE:
70 ORGANISM: Amsacta moorei entomopoxvirus
71 FEATURE:
72 NAME/KEY: CDS
73 LOCATION: complement (18..218)
74 FEATURE:

```

NAME/KEY: CDS
 LOCATION: complement (234..782)
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 852..1511
 US-08-107-755A-8

Query Match 3.5%; Score 40; DB 1; Length 1511;
 Best Local Similarity 52.4%; Pred. No. 0.065;
 Matches 88; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 941 ACCACTAGTCTTAAGAAGCTGTTTCTGTTTGGCGAAGGTGTATGTATGTTTAAAG 1000
 DB 453 AACCAATTAACCTTTTATTTATTTAGCCATTTTATCACAATAATGTTCTAAATCATTTTC 512
 QY 1001 TCAAAAATTTAGTAGGAATGCGTTACTAGTATACACGTGAAGTTCATTTGCAATGT 1060
 DB 513 TTCAAAAATTTAGCACTCTATGCGCAATTAATCATTAATTCATGATTTGATTTTC 572
 QY 1061 TTTAATAAATATGTTGCTTGTAGTTTAAAGTTGATTAATTTACTCT 1108
 DB 573 ATTAAATTAATTTATTTGTTTAAATGATTAATTAATTTCTTTATTTAAT 620

RESULT 6

US-08-544-332-8
 Sequence 8, Application US/08544332
 Patent No. 5935777
 GENERAL INFORMATION:
 APPLICANT: Moyer, Richard W.
 APPLICANT: Hall, Richard L.
 APPLICANT: Gruidl, Michael E.
 TITLE OF INVENTION: No. 5935777e1 Entomopoxvirus Expression System
 NUMBER OF SEQUENCES: 77
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Gerard H. Bencen
 STREET: 2421 N.W. 41st Street, Suite A-1
 CITY: Gainesville
 STATE: FL
 COUNTRY: USA
 ZIP: 32606
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/544,332
 FILING DATE:
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/991,867
 FILING DATE: 07-DEC-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/107,755
 FILING DATE: 19-AUG-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: WO 92/14818
 FILING DATE: 12-FEB-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/827,685
 FILING DATE: 30-JAN-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/657,584
 FILING DATE: 19-FEB-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: Bencen, Gerard H.
 REGISTRATION NUMBER: 35,746
 REFERENCE/DOCKET NUMBER: UF114.C4
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 904-375-8100
 TELEFAX: 904-372-5800

INFORMATION FOR SEQ ID NO: 8:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1511 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: unknown
 MOLECULE TYPE: DNA (genomic)
 ORIGINAL SOURCE:
 ORGANISM: Amsacta moorei entemopoxvirus
 FEATURE:
 NAME/KEY: CDS
 LOCATION: complement (18..218)
 FEATURE:
 NAME/KEY: CDS
 LOCATION: complement (234..782)
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 852..1511
 US-08-544-332-8

Query Match 3.5%; Score 40; DB 2; Length 1511;
 Best Local Similarity 52.4%; Pred. No. 0.065;
 Matches 88; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 941 ACCACTAGTCTTAAGAAGCTGTTTCTGTTTGGCGAAGGTGTATGTATGTTTAAAG 1000
 DB 453 AACCAATTAACCTTTTATTTATTTAGCCATTTTATCACAATAATGTTCTAAATCATTTTC 512
 QY 1001 TCAAAAATTTAGTAGGAATGCGTTACTAGTATACACGTGAAGTTCATTTGCAATGT 1060
 DB 513 TTCAAAAATTTAGCACTCTATGCGCAATTAATCATTAATTCATGATTTGATTTTC 572
 QY 1061 TTTAATAAATATGTTGCTTGTAGTTTAAAGTTGATTAATTTACTCT 1108
 DB 573 ATTAAATTAATTTATTTGTTTAAATGATTAATTAATTTCTTTATTTAAT 620

RESULT 7

US-08-852-629-11
 Sequence 11, Application US/08852629
 Patent No. 6106825
 GENERAL INFORMATION:
 APPLICANT: Moyer, Richard W
 APPLICANT: Li, Yi
 APPLICANT: Hall, Richard L
 TITLE OF INVENTION: ENTOMPOXYVIRUS-VERTEBRATE GENE DELIVERY
 NUMBER OF SEQUENCES: 17
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik
 STREET: 2421 N.W. 41st Street, Suite A-1
 CITY: Gainesville
 STATE: Florida
 COUNTRY: U.S.A.
 ZIP: 32606
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/852,629
 FILING DATE:
 CLASSIFICATION: 514
 ATTORNEY/AGENT INFORMATION:
 NAME: Bencen, Gerard H.
 REGISTRATION NUMBER: 35,746
 REFERENCE/DOCKET NUMBER: UF-184
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 407-426-7500
 TELEFAX: 407-839-8589
 INFORMATION FOR SEQ ID NO: 11:

SEQUENCE CHARACTERISTICS:
LENGTH: 4810 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-852-629-11

Query Match 3.5%; Score 40; DB 3; Length 4810;
Best Local Similarity 52.4%; Pred. No. 0.12;
Matches 88; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 941 ACCACTAGTCTCAAGAACTGTTCTGTTTGGCCGAAGTGTGATGTTTAC 1000
DB 459 AACCAATACCTTTTATTTATTTAGCCATTTTATCACAATAATGTTCTAAATCATTTTC 518
QY 1001 TCATAATATTTAGTGAAGAAATGCTTACTAGTATTAACCTGAAGTTCAATATGCAATGT 1060
DB 519 TTCAAAAATTTGACACTCATCTATGCCAATATATCATATTTATCTACGATATTGATTTTC 578
QY 1061 TTTAATAAATATTTGCTTGTGATTTAATTAAGTTGATATATCTCT 1108
DB 579 ATTAATTAATATTTGTTTATTTATGTATTAATATCTTTATTTAATAT 626

RESULT 8

US-08-852-629-15
Sequence 15, Application US/08852629
Patent No. 6106825

GENERAL INFORMATION:

APPLICANT: Moyer, Richard W
APPLICANT: LI, Yi
TITLE OF INVENTION: ENTOMOPOXYVIRUS-VERTEBRATE GENE DELIVERY
TITLE OF INVENTION: VECTOR AND METHOD
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: Florida
COUNTRY: U.S.A.
ZIP: 32606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/852,629
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Bence, Gerard H
REGISTRATION NUMBER: 35,746
REFERENCE/DOCKET NUMBER: UF-184
TELECOMMUNICATION INFORMATION:
TELEPHONE: 407-426-7500
TELEFAX: 407-839-8589
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 4838 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-852-629-15

Query Match 3.5%; Score 40; DB 3; Length 4838;
Best Local Similarity 52.4%; Pred. No. 0.12;
Matches 88; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 941 ACCACTAGTCTCAAGAACTGTTTCTGTTTGGCCGAAGTGTGATGTTTAC 1000
DB 459 AACCAATACCTTTTATTTATTTAGCCATTTTATCACAATAATGTTCTAAATCATTTTC 518
QY 1001 TCATAATATTTAGTGAAGAAATGCTTACTAGTATTAACCTGAAGTTCAATATGCAATGT 1060
DB 519 TTCAAAAATTTGACACTCATCTATGCCAATATATCATATTTATCTACGATATTGATTTTC 578
QY 1061 TTTAATAAATATTTGCTTGTGATTTAATTAAGTTGATATATCTCT 1108
DB 579 ATTAATTAATATTTGTTTATTTATGTATTAATATCTTTATTTAATAT 626

RESULT 9

US-08-232-463-14/C
Sequence 14, Application US/08232463
Patent No. 5670367

GENERAL INFORMATION:

APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOMLOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZ9pt-F15
US-08-232-463-14

Query Match 3.4%; Score 38.2; DB 1; Length 7218;
Best Local Similarity 6.6%; Pred. No. 0.48;
Matches 28; Conservative 207; Mismatches 190; Indels 0; Gaps 0;

QY 75 CAGGCTCAAGTGGGAGCGGACAGATGTTCTCTCAAGAGTGAACGGGTAGCCA 134

LENGTH: 1037 base pairs

TOPOLOGY: 1 linear

MOLECULE TYPE: DNA (genomic)

RESULT 11
 US-08-299-953-1
 : Sequence 1, Application US/08299953
 : Patent No. 5646333
 : GENERAL INFORMATION:
 : APPLICANT: Dobrys, Michael S. and Mandaci, Sevnur
 : TITLE OF INVENTION: A plant Promoter Useful for Directing the
 : TITLE OF INVENTION: Expression of Foreign Proteins to the Plant Epidermis
 : NUMBER OF SEQUENCES: 4
 : CORRESPONDENCE ADDRESS:
 : ADDRESS: Woodcock Washburn Kurtz Mackiewicz & No. 5646333ris
 : STREET: One Liberty Place 46th. Floor
 : CITY: Philadelphia
 : STATE: PA
 : ZIP: 19103
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: PatentIn Release #1.0, Version #1.25
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/299,953
 : FILING DATE: Herewith
 : CLASSIFICATION: 435
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Beardell, Lori Y.
 : REGISTRATION NUMBER: 34,293
 : REFERENCE/DOCKET NUMBER: NOVA-0003
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: 215-564-8960
 : TELEFAX: 215-568-3439
 : INFORMATION FOR SEQ ID NO: 1:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 2861 base pairs
 : TYPE: nucleic acid
 : STRANDEDNESS: double
 : TOPOLOGY: linear
 : MOLECULE TYPE: DNA (genomic)

Query Match	Score	DB	Length
3.28;	36;	DB 1;	Length 2861

Search completed: July 25, 2002, 19:07:32
Job time: 14143 sec

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 25, 2002, 18:05:39 ; Search time 2540.26 Seconds
(without alignments)
4006.166 Million cell updates/sec

Title: US-09-509-779-3

Perfect score: 754
Sequence: 1 ATGCGCAGCAGTGAAGACGG.....CAATTAAGTGCAGTTTAA 754

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_hic:*
13: em_gss_hum:*
14: em_gss_hum:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	689	91.4	793	9	AL547435
2	664.8	88.2	949	10	BG769106
3	661.6	87.7	797	10	B1832643
4	659.8	87.5	918	10	B1759082
5	649	86.1	822	10	B1668735
6	647.4	85.9	818	10	B1457840
7	644.4	85.5	805	10	BG708518
8	643	85.3	856	10	B1601855
9	640.2	84.9	754	10	BG766992
10	634.4	84.1	724	10	B1226556
11	624.6	82.8	855	10	BE747000
12	623.6	82.7	893	10	BG037017
13	616.2	81.7	797	10	B1858307
14	609	80.8	856	10	BG753323
15	608.8	80.7	767	10	B1601470
16	607.4	80.6	702	9	AI338342
17	606.6	80.5	847	10	BG773490

18	606	80.4	748	10	BG576768	BG576768	602599091
19	597.2	79.2	904	10	BG742338	BG742338	602631707
20	594.2	78.8	716	10	BG761806	BG761806	602718011
21	592.8	78.6	840	10	BG037022	BG037022	602287341
22	589.6	78.2	827	10	B1828930	B1828930	603075092
23	588.8	78.1	675	10	B1858784	B1858784	603386387
24	587.2	77.9	947	10	BG111792	BG111792	602285379
25	582.4	77.2	745	10	BG121625	BG121625	602351581
26	581.8	77.2	734	10	BG121625	BG121625	602351581
27	575.2	76.3	864	10	B1757881	B1757881	603030486
28	574.8	76.2	793	10	BG716023	BG716023	602677415
29	571.8	75.8	651	10	BG714665	BG714665	602677016
30	555.4	73.7	595	10	BF033587	BF033587	601453564
31	554.2	73.5	781	10	BG026092	BG026092	602292288
32	553.6	73.4	670	9	AM001261	AM001261	ww26907.x
33	553.4	73.4	793	10	BG111145	BG111145	602281555
34	546.2	72.4	715	10	BG718227	BG718227	602696194
35	532.2	70.3	696	10	BG715867	BG715867	602675284
36	530.2	70.3	600	10	B1668630	B1668630	603293943
37	515.4	68.4	600	9	AA521294	AA521294	aa79e09.s
38	509.2	67.5	673	10	BE398022	BE398022	601290582
39	488.4	64.8	565	10	BG054821	BG054821	nae90c06
40	486.4	64.5	596	10	BG054726	BG054726	nae88d06
41	485.8	64.4	604	9	AA521231	AA521231	aa79e09.s
42	479.4	63.6	598	10	BG181095	BG181095	602329223
43	473.8	62.8	561	9	AI769519	AI769519	wj24e03.x
44	473.4	62.8	549	9	AA583455	AA583455	na38c09.s
45	471.4	62.5	555	9	AA805121	AA805121	of49d09.s

ALIGNMENTS

RESULT 1
AL547435 793 bp mRNA linear EST 16-FEB-2001
LOCUS AL547435 LTI_NFL006.PL2 Homo sapiens cDNA clone CS0D1011YW21 5
DEFINITION AL547435 LTI_NFL006.PL2 Homo sapiens cDNA clone CS0D1011YW21 5
prime, mRNA sequence.
ACCESSION AL547435 GI:12881506
VERSION AL547435
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 793)
AUTHORS Li,W.B., Gruber,C., Jesse,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seque@genoscope.cns.fr, Web : www.genoscope.cns.fr.
Location/Qualifiers
1. 793
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0D1011YW21"
/clone_1ib="LTI_NFL006.PL2"
/tissue_type="placenta"
/note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@life.com
http://fulllength.invitrogen.com"

BASE COUNT 230 a 163 c 203 g 195 t 2 others
ORIGIN

[illegible]

OY	181	TGCTTACATGTCGAAGCTGAAAACCAACAAAGAGACGTGGTGTGCTGCGGGAATAGT	240
Db	209	TGCTTACATGTCGAAGCTGAAAACCAACAAAGAGACGTGGTGTGCTGCGGGAATAGT	268
OY	241	AATCATTCCTTCCACAACTGCTGCATGTCCCTGTGGGTGAAACACAAACAATCGCGCCCT	300
Db	269	AATCATTCCTTCCACAACTGCTGCATGTCCCTGTGGGTGAAACACAAACAATCGCGCCCT	328
OY	301	CTCTGCACAGCAGGACGTGGGTGGTCCAAAGAAATGCGCAAAATGAGAGTGGTTAGAGGCTTC	360
Db	329	CTCTGCACAGCAGGACGTGGGTGGTCCAAAGAAATGCGCAAAATGAGAGTGGTTAGAGGCTTC	388
OY	361	TTAGCGCAGTGTTCACAGACCCCTGGTGGATCTGTATATCCAGTGGCCATACAAAGGCTTGA	420
Db	389	TTAGCGCAGTGTTCACAGACCCCTGGTGGATCTGTATATCCAGTGGCCCTACAAAGGCTTGA	448
OY	421	ACACTACAGGGGATGAAATCTTCAAAATAGAGACCGATGGATCTGTGATC - "TTTGAGATC	478
Db	449	ACACTACAGTGGATGAAATCTTCAAAATAGAGACCGATGGATCTGTGTGATC "TTTGAGATC	508
OY	479	ATCAAAAGCCCTTGCTT - AGCATTTGTGCAGTTTATCTTCACAAATCTCTGTGATTAGA	536
Db	509	ATCAAAAGCCCTTGCTTACCATTTTGTCTAGTTTATCTTCAGAAATCTCTGTGATTAGA	568
OY	537	AGATATTTATTTAAAGGTGTCCTTCCATACCTCTGTGGTGTGTGCGGCACACAGCTTA	596
Db	569	AGATATTTATTTAAAGGTGTCCTTCCATACCTCTGTGGTGTGTGCGGCACACAGCTTA	628
OY	597	GAAAGTGTATTAATAAA - AAGAAAGAGCTCCAAATTAATACAC - TTATATATTATACCATT	654
Db	629	GAAAGTGTATTAATAAAACAGAAAGAGCTCCAAATTAATACACCTTATATATTATACCATT	688
OY	655	CTATACCAACAGGACGTGGAAGCAGTTTCAGAG - ACTTTTTCAGTGTATATGTTGATCACT	713
Db	689	CTATACCAACAGGACGTGGAAGCAGTTTCAGAGAACTTTTGCATGCTTATATGTTGATCACT	748
OY	714	TAAAAAAGATGTTACATATACAAATAAAGTCATTTAAA	754
Db	749	T - ACCAAGAAATGTTACATATACCAAT - AAGTCAGATTTAAA	787

RESULT	6
B1457840	
LOCUS	B1457840 818 bp mRNA
DEFINITION	603199212.F1 NIH_MGC_96 Homo sapiens CDNA clone IMAGE:5277723.5 ; linear EST 21-AUG-2001
ACCESSION	B1457840 B1457840
VERSION	B1457840.1 GI:15248496
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens

REFERENCE	1 (pages 1 to 818)
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ .
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished (1999)
COMMENT	Contact: Robert Strausberg, Ph.D.

Tissue Procurement: Miklos Palkovits, M.D., Ph.D.,
 CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shitarak
 Toshiyuki and Piero Carninci (RIKEN)
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: L14M11701 row: g column: 04
 High quality sequence stop: 780.

FEATURES	Location/Qualifiers
source	1. 818
	/organism="Homo sapiens"
	/db_xref="taxon:9606"

/clone="IMAGE:52777723"
 /clone_lib="NIH_MGC_96"
 /tissue_type="hypothalamus"
 /lab_host="DH10B"
 /note="Organ: brain; Vector: pBluescript (modified
 pBluescript KS+); Site.1: BamHI; Site.2: SalI-XhoI (gcgcagc
); Oligo-dn primed using primer 5'-TGTGTTTGTGTTTGTGTTVN-3',
 size-selected for average insert size 2.3 kb and
 normalized to ROP 5. This is a primary library enriched
 for full-length clones and constructed using the
 Cap-trapper method (Carninci, in preparation). Library
 constructed by M. Brownstein (NIH/NHRRI, National
 Institutes of Health). Note: this is a NIH_MGC Library."

Query Match	85.9%;	Score 647.4;	DB 10;	Length 818;
Best Local Similarity	95.6%;	Pred. No. 1.4e-168;		
Matches 733; Conservative	0;	Mismatches 21;	Indels 13;	Gaps 6;

QY	1	ATGGCCGACGTGGAAAGACGGAGAGAAACCTCGCCCTGAGCTCTACATCCGGGAGCTCA	60
Db	17	ATGGCCGACGTGGAAAGACGGAGAGAAACCTCGCCCTGAGCTCTACATCCGGGAGCTCA	76
QY	61	GCCCTCCAGTGGGAGGGACGAAGATGTTCTCCCTCAAGAACTGGACGCGGTGGCCATG	120
Db	77	GGCTCCAGTGGGAGGGAGCGACAAAGATGTTCTCCCTCAAGAACTGGACGCGGTGGCCATG	136
QY	121	TGGAGCTGGGACGTGGAGTGGGATACGTGCGGCATCTGCAGGGTCCAGGTATGATGCC	180
Db	137	TGGAGCTGGGACGTGGAGTGGGATACGTGCGGCATCTGCAGGGTCCAGGTATGATGTCG	196
QY	181	TGCTTTAGATGTCAAGCTGAAAAACAACAAAGAGACTGTGTGGTCTGGGAGAAATGT	240
Db	197	TGTCTTAGATGTCAAGCTGAAAAACAACAAAGAGACTGTGTGGTCTGGGAGAAATGT	256
QY	241	AATCATTCCTTCCACAACTGCTGCATGTCCCTGTGGGTGAAACAGAACAACTGTCGCC	300
Db	257	AATCATTCCTTCCACAACTGCTGCATGTCCCTGTGGGTGAAACAAACAACTGTCGCC	316
QY	301	CTCTGCCAGCAGACACTGGGTGGTCCAAAGAAATCGGCCAAATGAGAGTGGTTAAAGCTTC	360
Db	317	CTCTGCCAGCAGACACTGGGTGGTCCAAAGAAATCGGCCAAATGAGAGTGGTTAAAGCTTC	376
QY	361	TTAGGCGAGTGTTCAGAGCCCTGTGGATCTGTGAATCCATGCGCTCAAGAAGCTAGA	420
Db	377	TTAGGCGAGTGTTCAGAGCCCTGTGGATCTGTGAATCCATGCGCTCAAGAAGCTAGA	436
QY	421	ACACTACAGGGGATGAATTTCTTCAATAGAGCCGATGGATCTGTGTC -TTTGGACTC	478
Db	437	ACACTACAGGGGATGAATTTCTTCAATAGAGCCGATGGATCTGTGTCCTTTGGGACTC	496
QY	479	ATCAAAGCCCTGGTT - AGCATTTGTCAAGTTTATCTTCAGAAATTTCTGTGATTAGA	536
Db	497	ATCAAAGCCCTGGTTTACCATTTTGTCAAGTTTATCTTCAGAAATTTCTGTGATTAGA	556
QY	537	AGATATTTTAAAGTGTGCTCTCAACCTGTGTGGTGTGTCGGCACACAGCTTA	596
Db	557	AGATATTTTAAAGTGTGCTCTCTCAACCTGTGTGGTGTGTCGGCACACAGCTTA	616
QY	597	GAAGTGCTATAAAAAAGGAAAGAGCTCCAAATGGAATCAC -TTATTAATTTACCATTTTC	655
Db	617	GAAGTGCTATAAAAAAGGAAAGAGCTCCAAATGGAATCACCTTATTCATTTACCATTTTC	676
QY	656	TATAC -AACAGCGAGTGAAGCAGTTTC - -GAGACTTTTTCGATGCTTATGTTGATC	710
Db	677	TATACAAACAGCGAGTGAAGCAGTTTCAGAGAACTTTTTCGATGCTTATGTTGTTGATC	736
QY	711	AGTT - -AAAAAGATGTTACAGTAACAATTAAGTGCAGTTTAAA	754
Db	737	CAGTTAAAAAAGATGTTTACAGTAACAATTAAGTGCAGTTTAAA	783

```

RESULT 7
LOCUS Bg708518 805 bp mRNA linear EST 07-MAY-2001
DEFINITION 602670411F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4793091 5'
ACCESSION Bg708518
VERSION Bg708518.1 GI:13985940
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 805)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: Miklos Palokovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraaki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLM10672 row: f column: 04
High quality sequence stop: 740.
FEATURES
source
1..805
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="IMAGE:4793091"
/tissue_type="NIH_MGC_96"
/lab_host="DH10B"
/note="Organ: Brain; Vector: pBluescriptR (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTVA-3',
size-selected for average insert size 2.3 kb and
normalized to 10^5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT 214 a 183 c 214 g 194 t
ORIGIN
Query Match 85.5%; Score 644.4; DB 10; Length 805;
Best Local Similarity 94.6%; Pred. No. 9, 8e-168;
Matches 723; Conservative 0; Mismatches 31; Indels 10; Gaps 5;
QY 1 ATGGCGACGTGGAAGAGGAGAGAACTGGCCCTGCTCACTCCGGGAGCTCA 60
DB 26 ATGGCGACGTGGAAGAGGAGAGAACTGGCCCTGCTCACTCCGGGAGCTCA 85
QY 61 GGCCTCAAGTGGGAGGAGGAGCAAGATGTTCTCCCTCAAGAGTGAACGGCGTGCATG 120
DB 86 GGCCTCAAGTGGGAGGAGGAGCAAGATGTTCTCCCTCAAGAGTGAACGGCGTGCATG 145
QY 121 TGGAGCTGGAGCTGGAAGTCCGATACGTGGCCATCTGCAGGCTCAGAGTGAATGCC 180
DB 146 TGGAGCTGGAGCTGGAAGTCCGATACGTGGCCATCTGCAGGCTCAGAGTGAATGCC 205
QY 181 TGTCTAGATGTCAGCTGAAAAACAAGAGAGGAGCTGTGGCTGGGGAGATGT 240
DB 206 TGTCTAGATGTCAGCTGAAAAACAAGAGAGGAGCTGTGGCTGGGGAGATGT 265
QY 241 AATCATCTCTTCCCAACAATGCTGCATGCTCTGGTGGTGAACAGAAATCCTGCCCT 300
DB 266 AATCATCTCTTCCCAACAATGCTGCATGCTCTGGTGGTGAACAGAAATCCTGCCCT 325

```

```

QY 301 CTCTGCCAGGAGACTGGGTGCTCCAAAGATGGCAATGAGGTGTTAGAGGCTTC 360
DB 326 CTCTGCCAGGAGACTGGGTGCTCCAAAGATGGCAATGAGGTGTTAGAGGCTTC 385
QY 361 TTACGCGAGTTGTTCAAGAGCCCTGGTGCATCTTAAATCCAGTCCCTACAAAGGCTAGA 420
DB 386 TTACGCGAGTTGTTCAAGAGCCCTGGTGCATCTTAAATCCAGTCCCTACAAAGGCTAGA 445
QY 421 AACTACAGGGGATGAATTTCTCAATAGAGACCGGATGATCTGTGTC--TTTGGACTC 478
DB 446 AACTACAGGGGATGAATTTCTCAATAGAGACCGGATGATCTGTGTC--TTTGGACTC 505
QY 479 ATCAAGCCTTGGT--AGCATTTGCTAGTTTATCTTCGAATTCCTGTGATTTAGA 536
DB 506 ATCAAGCCTTGGT--AGCATTTGCTAGTTTATCTTCGAATTCCTGTGATTTAGA 565
QY 537 AGATTAATTTAATTA--AGTGGTCTCTTCTACCTCTGTGTGTGTGCGCACACACTT 595
DB 566 AGATTAATTTAATTAACAGGTGGTCTCTTCTACCTCTGTGTGTGTGCGCACACACTT 625
QY 596 AGAAGTCTTTAAAGAAAGAGAGCTCAATTTGATCACC--TTATATTTACCATTT 654
DB 626 AGAAGTCTTTAAAGAAAGAGAGCTCAATTTGATCACC--TTATATTTACCATTT 685
QY 655 CTATACACAGGAGAGTGGAGAGCAGTTTCGAG---ACTTTTGATGCTTATGTTGATC 710
DB 686 CTATACACAGGAGAGTGGAGAGCAGTTTCGAG---ACTTTTGATGCTTATGTTGATC 745
QY 711 AGTTAAATTAATGTTTACAGTAACAATTAATGTCAGTTTAA 754
DB 746 AGTTAAATTAATGTTTACAGTAACAATTAATGTCAGTTTAA 789

```

```

RESULT 8
LOCUS B1601855 856 bp mRNA linear EST 07-SEP-2001
DEFINITION B03244729F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5287014 5'
ACCESSION B1601855
VERSION B1601855.1 GI:15494794
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 856)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: Miklos Palokovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraaki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLM11725 row: j column: 07
High quality sequence stop: 820.
FEATURES
source
1..856
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="IMAGE:5287014"
/tissue_type="NIH_MGC_96"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescriptR (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTVA-3',

```

size-selected for average insert size 2.3 kb and normalized to R0T 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIH/NIHRI, National Institutes of Health). Note: this is a NIH_MGC Library.

BASE COUNT 227 a 186 c 227 g 216 t

ORIGIN

Query Match 85.3%; Score 643; DB 10; Length 856;
Best Local Similarity 95.3%; Pred. No. 2,4e-167;
Matches 729; Conservative 0; Mismatches 25; Indels 11; Gaps 6;

1 ATGGCCAGCTGGAGAGCGAGAGAAACCGCCCTGGCCTCTCACTCCGGAGCTCA 60
17 ATGGCCAGCTGGAGAGCGAGAGAAACCGCCCTGGCCTCTCACTCCGGAGCTCA 76
61 GGCTCCAGCTGGAGAGCGAGAGAAACCGCCCTGGCCTCTCACTCCGGAGCTCA 120
77 GGCTCCAGCTGGAGAGCGAGAGAAACCGCCCTGGCCTCTCACTCCGGAGCTCA 136
121 TGAAGCTGGAGAGCTGGAGAGCGAGAGAAACCGCCCTGGCCTCTCACTCCGGAGCTCA 180
137 TGAAGCTGGAGAGCTGGAGAGCGAGAGAAACCGCCCTGGCCTCTCACTCCGGAGCTCA 196
181 TGTCTTGAAGCTGGAGAGCGAGAGAAACCGCCCTGGCCTCTCACTCCGGAGCTCA 240
197 TGTCTTGAAGCTGGAGAGCGAGAGAAACCGCCCTGGCCTCTCACTCCGGAGCTCA 256
241 AATCATTCCTCCACACCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
257 AATCATTCCTCCACACCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 316
301 CTCTGCAGAGAGCTGGAGAGCGAGAGAAACCGCCCTGGCCTCTCACTCCGGAGCTCA 360
317 CTCTGCAGAGAGCTGGAGAGCGAGAGAAACCGCCCTGGCCTCTCACTCCGGAGCTCA 376
361 TTGAGCGAGTTGTCAGAGCGCTGGAGAGCGAGAGAAACCGCCCTGGCCTCTCACTCCGGAGCTCA 420
377 TTGAGCGAGTTGTCAGAGCGCTGGAGAGCGAGAGAAACCGCCCTGGCCTCTCACTCCGGAGCTCA 436
421 ACACTACAGGAGGATGATTTCTCAATATAGAGCGAGAGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 477
437 ACACTACAGGAGGATGATTTCTCAATATAGAGCGAGAGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 493
478 CATCAAAAGCCTGGT--AGCAATTTGCACTTTTATCTTCAAGAAATCTCTGTGATTAAG 535
497 CATCAAAAGCCTGGT--AGCAATTTGCACTTTTATCTTCAAGAAATCTCTGTGATTAAG 556
536 AAGTATATTTATTTAA--AGGTGCTCTCTCACTGCT 594
557 AAGTATATTTATTTAA--AGGTGCTCTCTCACTGCT 616
595 TAGAAGTCTATATAAAGAAAGAGCTCAATTTGAATCAAC--TTAATATTTACCATTT 653
617 TAGAAGTCTATATAAAGAAAGAGCTCAATTTGAATCAAC--TTAATATTTACCATTT 676
654 TCTATACACAGGAGCTGGAGAGCGAGAGAAACCGCCCTGGCCTCTCACTCCGGAGCTCA 709
677 TCTATACACAGGAGCTGGAGAGCGAGAGAAACCGCCCTGGCCTCTCACTCCGGAGCTCA 736
710 CAGTATAAAGAGATGTTACAGTAAATTAAGTGAATTAAG 754
737 CAGTATAAAGAGATGTTACAGTAAATTAAGTGAATTAAG 781

RESULT 9
BG766992 754 bp mRNA linear EST 15-MAY-2001
LOCUS 602740511F1 NIH_MGC_49 Homo sapiens cDNA clone IMAGE:4870251 5'
DEFINITION mRNA sequence.
ACCESSION BG766992

VERSION BG766992.1 GI:14077645
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 754)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
unpublished (1999)
JOURNAL

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: ARCC/DCID/DRP
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
Plate: LLCMI743 row: e column: 04
High quality sequence stop: 750.

FEATURES

source

1..754
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="IMAGE:4870251"
/clone_image="NIH_MGC_49"
/issue_type="melanotic melanoma, high MDR (cell line)"
/lab_host="DH10B (phage-resistant)"
/note="Organ: Skin; Vector: pOT8; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library." location/Qualifiers

BASE COUNT 205 a 160 c 197 g 192 t

ORIGIN

Query Match 84.9%; Score 640.2; DB 10; Length 754;
Best Local Similarity 96.9%; Pred. No. 1,4e-166;
Matches 728; Conservative 0; Mismatches 13; Indels 10; Gaps 7;

13 GAAGACGAGAGGAAGAACTGGCCCTGGCTCTCACTCCGGAGAGCTCAAGCTCAAGTCG 72
1 GAAGACGAGAGGAAGAACTGGCCCTGGCTCTCACTCCGGAGAGCTCAAGCTCAAGTCG 60
73 GGAGGCGACAAAGATGTTCTCCCTCAAGATGGAACCGCGGTGGCATGTGAGCTGGAC 132
61 GGAGGCGACAAAGATGTTCTCCCTCAAGATGGAACCGCGGTGGCATGTGAGCTGGAC 120
133 GTGAGAGCATACGCTGCGCATCTGAGAGGTCCAGAGTGAATGATGCTCTTATAGATG 192
121 GTGAGAGCATACGCTGCGCATCTGAGAGGTCCAGAGTGAATGATGCTCTTATAGATG 180
193 CAAGCTGAAGAAACAAAGAGAGAGCTGTTGTGCTGGGAGAGATTAATCATTTCTTC 252
181 CAAGCTGAAGAAACAAAGAGAGAGCTGTTGTGCTGGGAGAGATTAATCATTTCTTC 240
253 CACACCTGCTGCAATGCT 312
241 CACACCTGCTGCAATGCT 300
313 GACTGGGTGCTCCAAAGATCGCAATGAGAGTGTAGAGGCTCTTAGCGCACTTG 372
301 GACTGGGTGCTCCAAAGATCGCAATGAGAGTGTAGAGGCTCTTAGCGCACTTG 360
373 TTGAGAGCCCTGGTGGATCTTTATATCCAGTCCCTCAAGAGCTTAGACACTAGAGGG 432
361 TTGAGAGCCCTGGTGGATCTTTATATCCAGTCCCTCAAGAGCTTAGACACTAGAGGG 420

Query Match	84.1%	Score 634.4	DB 10	Length 724
Best Local Similarity	96.6%	Pred. No. 5	6e-165	
Matches 705	Conservative 0	Mismatches 11	Indels 14	Gaps 5
QY 18	CGGAGAGGAACCTGCGCCCTGCGCTCTCACTCCGGAGGCTCAGGCTCCAGTGGGAGG	77		
Db 2	CGGAGAGGAACCTGCGCCCTGCGCTCTCACTCCGGAGGCTCAGGCTCCAGTGGGAGG	61		
QY 78	CGAGCAAGATTGTCCTCCCTCAAGAAGTGAAGCGGGGTCGATGGAGCTGGGACCTGGA	137		
Db 62	CGAGCAAGATTGTCCTCCCTCAAGAAGTGAAGCGGGGTCGATGGAGCTGGGACCTGGA	121		
QY 138	GTGCGATACGTGCGCCATCTGCAAGGTCGACGATGATGATGATGATGATGATGATGATG	197		
Db 122	GTGCGATACGTGCGCCATCTGCAAGGTCGACGATGATGATGATGATGATGATGATGATG	174		
QY 198	TGAAACAAACAGAGAGAGCTGTGTGGGCTGGGGAGATGTATATCATCTCTCCACAA	257		
Db 175	TGAAACAAACAGAGAGAGCTGTGTGTGGGCTGGGGAGATGTATATCATCTCTCCACAA	234		
QY 258	CTGCTGCATGTCCCTGTGGGTGAACAGAACATCGCTGCTCTGCGACGAGAGCTG	317		
Db 235	CTGCTGCATGTCCCTGTGGGTGAACAGAACATCGCTGCTCTGCGACGAGAGCTG	294		
QY 318	GGTGGTCCAAAGATGGGAAATGACAGTGGTTAGAAAGCTTCTTGGCAGATGTTCAAG	377		
Db 295	GGTGGTCCAAAGATGGGAAATGACAGTGGTTAGAAAGCTTCTTGGCAGATGTTCAAG	354		
QY 378	AGCCCTGTGATCTTTGTAATCCAGTGCCTCAAAAGGCTAGAACACTACAGGGATGAA	437		
Db 355	AGCCCTGTGATCTTTGTAATCCAGTGCCTCAAAAGGCTAGAACACTACAGGGATGAA	414		
QY 438	TTCCTCAATAGAGAGCGGATGATGATGATGATGATGATGATGATGATGATGATGATG	493		
Db 415	TTCCTCAATAGAGAGCGGATGATGATGATGATGATGATGATGATGATGATGATGATG	474		
QY 494	AGCATTTGTCAATTTATCTTCAGAAATTCCTGTGATTTAGAGATTAATTTATTTAAAG	553		
Db 475	AGCATTTGTCAATTTATCTTCAGAAATTCCTGTGATTTAGAGATTAATTTATTTAAAG	534		
QY 554	TGTCCTTCCTTCACCTGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	613		
Db 535	TGTCCTTCCTTCACCTGTGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	594		
QY 614	GAAAGAGCTCCAAATTTGATCACAC-TTATTAATTAATTAATTAATTAATTAATTAATTA	672		
Db 595	GAAAGAGCTCCAAATTTGATCACAC-TTATTAATTAATTAATTAATTAATTAATTAATTA	654		
QY 673	AAGCAGTTTC-GAGACTTTTTCGATGCTTATGTTGATCACTTAAAAAGATTTTACA	730		
Db 655	AAGCAGTTTCGAGAACTTTTTCGATGCTTATGTTGATCACTTAAAAAGATTTTACA	714		
QY 731	GTAAACAATA 740			
Db 715	GTAAACAATA 724			
RESULT 11				
BE747000	855 bp	mRNA	linear	EST 15-SEP-2000
LOCUS	60158074.F1	NIH_MGC_9	Homo sapiens	cdna clone IMAGE:3929459 5'
DEFINITION	mRNA sequence.			
ACCESSION	BE747000			
VERSION	BE747000.1	GI:10160992		
KEYWORDS	EST.			
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryotes; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
ATTENDS	1 (bases 1 to 855)			

COMMENT

Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-remail.nih.gov
 Tissue Procurement: DCTD/DNP
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed By: The I.M.A.G.E. Consortium (LNLN)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov,
 Plate: L1CM762 row: e column: 12
 High quality sequence stop: 767.
 Location/Qualifiers

FEATURES

source

```

1..855
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3929459"
/clone_lib="NIH_MGC_9"
/tissue_type="adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: ovary; Vector: pOTB7; Site: 1; XhoI; Site: 2;
ECORI; cDNA made by oligo-dT priming. Directionally
cloned into EORI/XhoI sites using the following 5'
adapter: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
```

BASE COUNT 250 a 161 c 238 g 205 t 1 others

ORIGIN

Query Match 82.8%; Score 624.6; DB 10; Length 855;
 Best Local Similarity 97.6%; Pred. No. 3.1e-162;
 Matches 698; Conservative 0; Mismatches 9; Indels 8; Gaps 6;

```

QY 30 CTGCGCCCTGCGCTCTCACTCCGGAGCTCAGGCTCCAAAGTGGGAGCGGACAGATGTT 89
    |||||||
DB 1 CTGCGCCCTGCGCTCTCACTCCGGAGCTCAGGCTCCAAAGTGGGAGCGGACAGATGTT 60
    |||||||
QY 90 CTCCTCAAGAGTGAAGCGGCTGCGCATGTGAGCTGGGAGCTGGAGTGCATACGTC 149
    |||||||
DB 61 CTCCTCAAGAGTGAAGCGGCTGCGCATGTGAGCTGGGAGCTGGAGTGCATACGTC 120
    |||||||
QY 150 CGCATCTCAGGCTCCAGAGTATGATGCTCTCTTAAAGTCAAGCTGAACAAACA 209
    |||||||
DB 121 CGCATCTCAGGCTCCAGAGTATGATGCTCTCTTAAAGTCAAGCTGAACAAACA 180
    |||||||
QY 210 AGAGCACTGTGTGTGTGGGAGATGTAATCATCTCTCCACACTGCTGCATGTC 269
    |||||||
DB 181 AGAGCACTGTGTGTGTGGGAGATGTAATCATCTCTCCACACTGCTGCATGTC 240
    |||||||
QY 270 CCTGTGGTGAAGAAACAATCGCGCTCTCGCCAGCAAGACTGGTGGTCCAAAG 329
    |||||||
DB 241 CCTGTGGTGAAGAAACAATCGCGCTCTCGCCAGCAAGACTGGTGGTCCAAAG 300
    |||||||
QY 330 AATCGCAATGAGATGTTAGAGGCTTCTAGCGCAGTTGTTCAAGCCCTGTGGA 389
    |||||||
DB 301 AATCGCAATGAGATGTTAGAGGCTTCTAGCGCAGTTGTTCAAGCCCTGTGGA 360
    |||||||
QY 390 TCTTGAATCCAGTCCCTACAAAGCTAGAACTACAGGGAGATCTTCAAAATG 449
    |||||||
DB 361 TCTTGAATCCAGTCCCTACAAAGCTAGAACTACAGGGAGATCTTCAAAATG 420
    |||||||
QY 450 GAGCCGATGATCTGTGTC--TTTGGACTCATCAAGCCTTGG--TTACACTTTGCACT 506
    |||||||
DB 421 GAGCCGATGATCTGTGTC--TTTGGACTCATCAAGCCTTGG--TTACACTTTGCACT 480
    |||||||
QY 507 TTTATCTTCAAGAAATCTCTGTATAGAAATATTTATTTAAAGTGGTCTTCTTAC 566
    |||||||
DB 481 TTTATCTTCAAGAAATCTCTGTATAGAAATATTTATTTAAAGTGGTCTTCTTAC 540
    |||||||
QY 567 CTCTGTGTGTGTGTGGGACACACAGCTTGAAGTCTTATAAAAAGAAAGAGCTCCA 626
    |||||||
DB 541 CTCTGTGTGTGTGTGGGACACACAGCTTGAAGTCTTATAAAAAGAAAGAGCTCCA 600
    |||||||
```

```

QY 627 ATTGAATCAC--TTATTAATTACCATTTCTATACAAC--GGCAGTGGAGAGTTTC-- 682
    |||||||
DB 601 ATTGAATCACCTTTATTAATTACCATTTCTATACAACGGGAGAGTGGAGAGTTTACG 660
    |||||||
QY 683 GAGACTTTTTCGATGCTTATGTTGATGATGATGATGATGATGATGATGATGATGAT 737
    |||||||
DB 661 AAGACTTTTTCGATGCTTATGTTGATGATGATGATGATGATGATGATGATGATGAT 714
    |||||||
```

RESULT 12

LOCUS

BG037017 893 bp mRNA linear EST 24-JAN-2001

DEFINITION

602287310P1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4374483 5',

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-remail.nih.gov
 Tissue Procurement: Miklos Palakovits, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiroki
 Toshiyuki and Piero Carninci (RIKEN)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: L1AM10038 row: d column: 04
 High quality sequence stop: 741.
 Location/Qualifiers

FEATURES

source

```

1..893
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4374483"
/clone_lib="NIH_MGC_96"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/notes="Organ: Brain; Vector: pBluescript (modified
pBluescript KS+); Site: 1; BamHI; Site: 2; SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',
size-selected for average insert size 2.3 kb and
normalized to R0.5. This is a primary library enriched
for full-length clones and constructed using the
cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."
```

BASE COUNT

ORIGIN

Query Match 82.7%; Score 623.6; DB 10; Length 893;
 Best Local Similarity 97.1%; Pred. No. 6e-162;
 Matches 678; Conservative 0; Mismatches 14; Indels 6; Gaps 4;

```

QY 1 ATGGCCGAGCTGGAAGAGGAGGAGAAACCTCGCCCTGCTCATCCGGAGCTCA 60
    |||||||
DB 29 ATGGCCGAGCTGGAAGAGGAGGAGAAACCTCGCCCTGCTCATCCGGAGCTCA 88
    |||||||
QY 61 GGCCTCAAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 120
    |||||||
DB 89 GGCCTCAAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 148
    |||||||
QY 121 TGGAGCTGGAGCTGAGTACGTGCGCATCTCGAGGCTCAGGTGATGATGCC 180
    |||||||
```

```

Db 149 TGGAGCTGGAGCTGAGTGCATACGTGCGCCATCTCGAGGCTCCAGGTGATGATGCC 208
Oy 181 TGTCTTAATGTCAAGCTGAAGAAACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
Db 209 TGTCTTAATGTCAAGCTGAAGAAACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 268
Oy 241 ATTCATTCCTTCCAACTGCTGCATGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
Db 269 AATCATTCCTTCCAACTGCTGCATGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 328
Oy 301 CTCTCCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
Db 329 CTCTCCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 388
Oy 361 TTAGCGCATGTTTCAGAGCCCTGCTGATCTTGAATCCAGTCCCTACAAAGGCTAGA 420
Db 389 TTAGCGCATGTTTCAGAGCCCTGCTGATCTTGAATCCAGTCCCTACAAAGGCTAGA 448
Oy 421 ACACATACAGGGAGATGAATCTTCAAAATAGAGCCGATGATCTGTGCTC--TTTGACTC 478
Db 449 ACACATACAGGGAGATGAATCTTCAAAATAGAGCCGATGATCTGTGCTC--TTTGACTC 508
Oy 479 ATCAAGGCTTGTGTT--AGCATTTGTGCTGTTTAACTTCAAGAAATCTCTGTGATAGA 536
Db 509 ATCAAGGCTTGTGTTAGCATTTGTGCTGTTTAACTTCAAGAAATCTCTGTGATAGA 568
Oy 537 AGATTAATTAATTA--AGGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 595
Db 569 AGATTAATTAATTAACAGGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 628
Oy 596 AGAAGTGTATTAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 654
Db 629 AGAAGTGTATTAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 688
Oy 655 CTATACACAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 692
Db 689 CTATACACAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 726

RESULT 13
Bi858307 797 bp mRNA linear EST 10-OCT-2001
LOCUS 603384080F1 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:5393131 5'
DEFINITION mRNA sequence.
ACCESSION Bi858307
VERSION Bi858307.1 GI:15999054
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
            Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
            NIH-MGC http://mgc.nci.nih.gov/
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
            Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-remail.nih.gov
            Tissue Procurement: DCTD/DIR
            CDNA Library Preparation: Life Technologies, Inc.
            DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: L1AM12001 row: 0 column: 20
            High quality sequence stop: 718.
            Location/Qualifiers
                1..797
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="IMAGE:5393131"
                /clone_11b="NIH_MGC_87"
                /tissue_type="mammary adenocarcinoma, cell line"

```

```

/lab_host="dh10b (phage-resistant)"
/note="Organ: breast; Vector: pCMV-SPORT6; Site_1: NotI;
      Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
      Average insert size 1.383 kb. Library enriched for
      full-length clones and constructed by Life Technologies.
      Note: this is a NIH-MGC Library."
BASE COUNT      229 a      166 c      204 g      198 t
ORIGIN

Query Match      81.7%; Score 616.2; DB 10; Length 797;
Best Local Similarity 95.0%; Pred. No. 6,5e-160;
Matches 725; Conservative 0; Mismatches 23; Indels 15; Gaps 8;

Oy 4 GCCACGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 63
Db 1 GCCACGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60
Oy 64 TCCAAAGTCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 123
Db 61 TCCAAAGTCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
Oy 124 AGCTGGAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 183
Db 121 AGCTGGAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
Oy 184 CTTAGATGTCAAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 243
Db 181 CTTAGATGTCAAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
Oy 244 CATTCCTTCCCAACATGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 303
Db 241 CATTCCTTCCCAACATGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
Oy 304 TGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 363
Db 301 TGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
Oy 364 GCGCAGTGTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 423
Db 361 GCGCAGTGTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
Oy 424 CTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 481
Db 421 CTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
Oy 482 AAGCCTTGCTT--AGCATTTGTGATTTATCTTCAAGAAATCTCTGCGATTAGAGAGA 539
Db 481 AAGCCTTGCTTATGATTTGTGATTTATCTTCAAGAAATCTCTGCGATTAGAGAGA 540
Oy 540 TAAATTAAT--AAAGTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 598
Db 541 TAAATTAATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
Oy 599 AGTGTATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 656
Db 601 AGTGTATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
Oy 657 ATACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 714
Db 661 AT--ACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 717
Oy 715 AAA--AAAGATGTTACAGTAACAAGAAATTAAGTGAAGTTTAA 754
Db 718 AAACACAGAGATTTTACAGTTTACCAATTAAGTGAAGTTTAA 760

RESULT 14
Bi753323 856 bp mRNA linear EST 15-MAY-2001
LOCUS 60273140F1 NIH_MGC_43 Homo sapiens cDNA clone IMAGE:4875489 5'
DEFINITION mRNA sequence.
ACCESSION Bi753323

```

```

VERSION      BG753323.1  GI:14063976
KEYWORDS     EST.
SOURCE       human.
ORGANISM     Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE    1 (bases 1 to 856)
             NIH-MGC http://mhc.nci.nih.gov/.
AUTHORS     National Institutes of Health, Mammalian Gene Collection (MGC).
JOURNAL      Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
             Email: cga@bbs-r@mail.nih.gov
             Tissue Procurement: ATCC
             CDNA Library Preparation: Ling Hong/Rubin Laboratory
             DNA Sequencing by: Incyte Genomics, Inc.
             Cloned through the I.M.A.G.E. Consortium/LLNL at:
             http://image.llnl.gov
             Plate: LLMC1756 row: 0 column: 10
             High quality sequence stop: 767.
FEATURES
  source
    1..856
       /organism="Homo sapiens"
       /db_xref="taxon:9606"
       /clone="IMAGE:4875489"
       /clone_lib="NIH_MGC_43"
       /tissue_type="normal pigmented retinal epithelium".
       /lab_host="DH10B (phage-resistant)".
       /note="Organ: eye; Vector: pOTR7; Site_1: XhoI; Site_2:
       EcoRI; CDNA made by oligo-dT priming. Directionally
       cloned into EcoRI/XhoI sites using the following 5'
       adaptor: GGCACGAG(G). Library constructed by Ling Hong
       in the laboratory of Gerald M. Rubin (University of
       California, Berkeley) using ZAP-CDNA Synthesis Kit
       (Stratagene) and Superscript II RT (Life Technologies).
       Note: this is a NIH-MGC Library. 1"
BASE COUNT   231 a      214 c      211 g      200 t
ORIGIN
Query Match      80.8%; Score 609; DB 10; Length 856;
Best Local Similarity 94.1%; Pred. No. 6,6e-158;
Matches 712; Conservative 0; Mismatches 30; Indels 15; Gaps 7;
QY 1 ATGGCCGAGCTGGAAGAGAGAGAACTGGCCCTCTCTACTCCGGAGCTCA 60
    |||||||
DB 18 ATGGCCGAGCTGGAAGAGAGAGAACTGGCCCTCTCTACTCCGGAGCTCA 77
QY 61 GGCTCCAGTCGGGAGGCGAAGATGTTCTCCCTCAAGAAAGTGAACGGGGGCATG 120
    |||||||
DB 78 GGCTCCAGTCGGGAGGCGAAGATGTTCTCCCTCAAGAAAGTGAACGGGGGCATG 137
QY 121 TGGAGCTGGAGCTGAGTGCATAGTGGCCATCTGACGGGTCCAGTGTATGCC 180
    |||||||
DB 138 TGGAGCTGGAGCTGAGTGCATAGTGGCCATCTGACGGGTCCAGTGTATGCC 197
QY 181 TGTCTTAGATGTCAAGCTGAAAACAAGAGAGACTGTGTGTCTGGGAGATGT 240
    |||||||
DB 198 TGTCTTAGATGTCAAGCTGAAAACAAGAGAGACTGTGTGTCTGGGAGATGT 257
QY 241 AATCATTTCTTCCACAACCTGCTCATGTCCCTGGTGAACAACAATCGCTGCC 300
    |||||||
DB 258 AATCATTTCTTCCACAACCTGCTCATGTCCCTGGTGAACAACAATCGCTGCC 317
QY 301 CTCCTGCAGCAGACTGTGGTGTCCAAAGATCGCAATAGAGATGTTAGAGCTTC 360
    |||||||
DB 318 CTCCTGCAGCAGACTGTGGTGTCCAAAGATCGCAATAGAGATGTTAGAGCTTC 377
QY 361 TTAGCCGAGTTGTTCAGAGCCCTGTGTGATCTTGTATTCAGTCCCTACAAAGCTAGA 420
    |||||||
DB 378 TTAGCCGAGTTGTTCAGAGCCCTGTGTGATCTTGTATTCAGTCCCTACAAAGCTAGA 437

```

```

QY 421 ACACACAGGGGATGAATCTTCAATAGAGCCGATGATCTGTGTC- -TTTGACATC 478
    |||||||
DB 438 ACACACAGGGGATGAATCTTCAATAGAGCCGATGATCTGTGTC- -TTTGACATC 497
QY 479 ATCAAAAGCCCTGTT- -AGCATTTGTCAGTTTATCTTCAAAATCTCTGTGATTAAC 536
    |||||||
DB 498 ATCAAAAGCCCTGTTGATTTGTCAGTTTATCTTCAAAATCTCTGTGATTAAC 557
QY 537 AGATATTTATTTAA-AGGTGTCCTTCCATCCCTGCTGGTGTGTCGACACAGCTT 595
    |||||||
DB 558 AGATATTTATTTAA-AGGTGTCCTTCCATCCCTGCTGGTGTGTCGACACAGCTT 617
QY 596 AGAAGTG-CTATTAAGAAAGAGAGAGCTCCAAATGAATCAC- -TTATATTTACCAT 653
    |||||||
DB 618 AGAAGTGCGATTAAGAAAGAGAGAGCTCCAAATGAATCAC- -TTATATTTACCAT 677
QY 654 TCTATCAACAGCAGCTGGAAGCAGTTTGAGACTTTTCGATG- -TTATGTTG 707
    |||||||
DB 678 TCTATCAACAGCAGCTGGAAGCAGTTTGAGACTTTTCGATG- -TTATGTTG 737
QY 708 ATCAGTTAA- -AAAGATGTTACAGTAAACAATAA 742
    |||||||
DB 738 ATCAGTTAA- -AAAGATGTTACAGTAAACAATAA 774
RESULT 15
BI601470 767 bp mRNA linear EST 07-SEP-2001
LOCUS 603249053F1 NIH_MGC_96 Homo sapiens CDNA clone IMAGE:5300845 5',
DEFINITION mRNA sequence.
ACCESSION BI601470
VERSION BI601470.1 GI:15494409
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 767)
          NIH-MGC http://mhc.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cga@bbs-r@mail.nih.gov
          Tissue Procurement: Miklos Palovits, M.D., Ph.D.
          CDNA Library Preparation: Michael J. Brownstein (NHRH), Shiraki
          Toshiyuki and Piero Carninci (RIKEN)
          CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
          DNA Sequencing by: Incyte Genomics, Inc.
          Cloned through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: LLMC1761 row: 1 column: 14
          High quality sequence stop: 701.
FEATURES
  source
    1..767
       /organism="Homo sapiens"
       /db_xref="taxon:9606"
       /clone="IMAGE:5300845"
       /clone_lib="NIH_MGC_96"
       /tissue_type="hypothalamus"
       /lab_host="DH10B"
       /note="Organ: brain; Vector: pBluescript (modified
       pluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
       ); Oligo-dT primed using primer 5'-TTTATTTTATTTTATTT-3',
       size-selected for average insert size 2.3 kb and
       normalized for R05. This is a primary library enriched
       for full-length clones and constructed using the
       Cap-trapper method (Carninci, in preparation) library
       constructed by M. Brownstein (NIMH/NHRH, National
       Institutes of Health). Note: this is a NIH-MGC Library."
BASE COUNT   192 a      173 c      214 g      188 t
ORIGIN

```


Query Match	80.7%;	Score 608.8;	DB 10;	Length 767;
Best Local Similarity	94.2%;	Pred. No. 7.2e-158;		
Matches 688;	Conservative	0;	Mismatches 32;	Indels 10;
				Gaps 5;

OY	1	ATGCGCAGCTGTGAAGACGGAGAGAAACCTGGCCCTGGCCCTCTACCTCGGGAGGCTCA	60
Db	35	ATGGCCACAGCTGGAAAGACGAGAGAAACCTGGCCCTGGCCCTCTACCTCGGGAGGCTCA	94
OY	61	GGCTCCAAAGTCGGGAGGCGACAGATGTTCTCCCTCAAGAAGGGAACCGGTGGCCATG	120
Db	95	GGCTCCAAAGTCGGGAGGCGACAGATGTTCTCCCTCAAGAAGGGAACCGGTGGCCATG	154
OY	121	TGGAGCTGGGACGTGGAGTSCGATACGTGGCCATCTGCAGGGTCCAGGTGATGATGCC	180
Db	155	TGGAGCTGGGACGTGGAGTSCGATACGTGGCCATCTGCAGGGTCCAGGTGATGATGCC	214
OY	181	TGCTCTAGATGTCAAGCTGAAACCAACAGAGAGACTGTGTTGTGGTCTGGGGAGAGTGT	240
Db	215	TGCTCTAGATGTCAAGCTGAAACCAACAGAGAGACTGTGTTGTGGTCTGGGGAGAGTGT	274
OY	241	AATCATTCCTTCCACAACCTGTGCATGTCCCTGGGTGGGAAACAGAAACATCGTCCTCT	300
Db	275	AATCATTCCTTCCACAACCTGTGCATGTCCCTGGGTGGGAAACAGAAACATCGTCCTCT	334
OY	301	CTCTGCCAGCAGGACTGGGTGGTCCAAAGAATCGGCAAAATGAGAGTGCTTGAAGGCTTC	360
Db	335	CTCTGCCAGCAGGACTGGGTGGTCCAAAGAATCGGCAAAATGAGAGTGCTTGAAGGCTTC	394
OY	361	TTAGCGCAGTGTTCACAGACCCTGGTGGATCTGTATCCAGTCCCTCAAAAGGCTAGA	420
Db	395	TTAGCGCAGTGTTCACAGACCCTGGTGGATCTGTATCCAGTCCCTCAAAAGGCTAGA	454
OY	421	ACACACAGGGGAGTAATCTTCAATAGGAGCGCATGATCTGTGGTC--TTTGGACTC	478
Db	455	ACACACAGGGGAGTAATCTTCAATAGGAGCGCATGATCTGTGGTC--TTTGGACTC	514
OY	479	ATCAAAAGCCTTGGTT--AGCATTTGTCAGTTTATCTTCAGAAATTCCTGTGATTTAAG	536
Db	515	ATCAAAAGCCTTGGTTTATGACATTTTGTGCAGTTTATCTTCAGAAATTCCTGTGATTTAAG	574
OY	537	ACATATATTTATTTAA--GGTGGTCCCTCCACCTCGTGGGTGGTGGCCGACACAGCTT	595
Db	575	ACATATATTTATTTAA--GGTGGTCCCTCCACCTCGTGGGTGGTGGCCGACACAGCTT	634
OY	596	AGAAATGCTATTTAAAGGAAGAAGAGCTCCAAATTTGAATCACTATATATTTACCATTTTC	655
Db	635	AGAAATGCTATTTAAAGGAAGAAGAGCTCCAAATTTGAATCACTATATATTTACCATTTTC	694
OY	656	TA--TACAAACAGCAGTGAAGCAATTTGCA--GACTTTTTCAGTCTTATGTTGATG	710
Db	695	TAATACAAACAGCAGTGAAGCAAGCTTCAGAGAACTTTTGCATCTTATGTTGATG	754
OY	711	AGTTAAAAA 720	
Db	755	ACGTTTAAAA 764	

Search completed: July 25, 2002, 18:05:48
Job time: 14227 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 25, 2002, 14:08:41 ; Search time 2540.26 Seconds
(without alignments)
6057.068 Million cell updates/sec

Title: US-09-509-779-1

Perfect score: 1140

Sequence: 1 GTTCTGGCGCCGCCATGCG.....ACTAATTCATCAATTAAATG 1140

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues 27472414

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Database :

EST:*
1: em_estbba:*
2: em_estbhu:*
3: em_estlin:*
4: em_estm:*
5: em_estcov:*
6: em_estcpl:*
7: em_estro:*
8: em_hic:*
9: gp_est1:*
10: gp_est2:*
11: gp_hic:*
12: gp_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1069.4	93.8	1132	11 AK003963	AK003963 Mus muscu
2	1041	91.3	1114	11 AK011328	AK011328 Mus muscu
3	1032.4	90.6	1084	11 AK003248	AK003248 Mus muscu
4	1004.4	88.1	1078	11 AK002628	AK002628 Mus muscu
5	765	67.1	820	10 B1732003	B1732003 603355869
6	753.4	66.1	1079	10 B1410936	B1410936 602962128
7	700.8	61.5	722	11 AK007588	AK007588 Mus muscu
8	686.6	60.2	731	10 B1647378	B1647378 603279634
9	673.4	59.1	978	10 B1156184	B1156184 602903365
10	665.4	58.4	916	10 B6917948	B6917948 602820877
11	651.4	57.1	754	10 B1147544	B1147544 602913971
12	651	57.1	704	10 BF018804	BF018804 ux83b03.x
13	649.2	56.9	946	10 BF383735	BF383735 602044644
14	646.2	56.7	737	10 B1556918	B1556918 603339970
15	624.4	54.8	822	10 BF302267	BF302267 602031233
16	612.2	53.7	755	10 BF579825	BF579825 602095851
17	597.2	52.4	620	10 BE628111	BE628111 uu29a11.x

c	18	596	52.3	651	9	AM555800
	19	575	50.4	575	10	BF023423
	20	568.8	49.9	735	10	B1153818
	21	567	49.7	568	10	BF720947
	22	556.8	48.8	577	9	AA230335
	23	554.8	48.7	574	10	B6087875
	24	547.2	48.0	576	9	AA221462
	25	542.2	47.6	760	10	BF163003
	26	542	47.5	558	9	AM542025
	27	537.4	47.1	852	9	BF159848
	28	536.4	47.1	627	11	AK006703
	29	530.6	46.5	847	10	BF140544
	30	525.6	46.1	881	10	BE915897
c	31	522.4	45.8	589	10	BC228832
	32	522	45.8	522	10	BE372989
	33	519	45.5	520	9	AM989997
	34	518.6	45.5	687	10	CB8520
	35	513	45.0	562	10	BE824793
c	36	512.4	44.9	563	10	BG143878
	37	512	44.9	513	10	BE850994
c	38	505.4	44.3	538	9	AM491099
	39	504.8	44.3	922	10	BF161048
	40	502	44.0	513	9	AA032932
	41	501.8	44.0	542	9	AA475269
	42	501	43.9	501	10	BE628851
	43	494	43.3	494	10	BC277865
c	44	493	43.2	555	10	BF730110
	45	491	43.1	507	10	BM022893

ALIGNMENTS

RESULT 1	AK003963	1132 bp	mus musculus 18 days embryo whole body cDNA, RIKEN full-length enriched library, clone:1110029M05:ring finger protein 7, full insert sequence.
LOCUS	AK003963		
DEFINITION	AK003963.1 GI:12834939		
ACCESSION	AK003963		
VERSION	HTC; CAP trapper.		
KEYWORDS	Mus musculus (strain:C57BL/6J) 18 days embryo cDNA to mRNA, clone.lib:RIKEN full-length enriched mouse cDNA library		
SOURCE	clone:1110029M05.		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	Carninci, P. and Hayashizaki, Y.		
TITLE	High-efficiency full-length cDNA cloning		
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)		
MEDLINE	99279253		
PUBMED	10349636		
REFERENCE	2 (sites)		
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.		
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes		
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)		
MEDLINE	20499374		
PUBMED	11042159		
REFERENCE	3 (sites)		
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuana, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kasahara, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.		
TITLE	RIKEN integrated sequence analysis (RISA) system--384-Format		
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)		

RESULT	2
LOCUS	AK011328
DEFINITION	Mus musculus 10 days embryo whole body cDNA, RIKEN full-length enriched library, clone:2610005N22::ing finger protein 7, full insert sequence.
ACCESSION	AK011328
VERSION	AK011328.1 GI:12847380
KEYWORDS	HTC; CAP trapper.
SOURCE	Mus musculus (strain:C57BL/6J) 10 days embryo CDNA to mRNA, clone_11b:RIKEN full-length enriched mouse CDNA library clone:2610005N22.
ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS	Carninci,P. and Hayashizaki,Y. 1 (sites)
TITLE	High-efficiency full-length cDNA cloning
JOURNAL MEDLINE	Meth. Enzymol. 303, 19-44 (1999) 99279253
PUBMED	10349636
REFERENCE	2 (sites)
AUTHORS	Garnincil,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,T., Muramatsu,M. and Hayashizaki,Y.
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL MEDLINE	Genome Res. 10 (10), 1617-1630 (2000) 20499374
PUBMED	11042159
REFERENCE	3 (sites)
AUTHORS	Shibata,K., Itoh,M., Aizawa,K., Nagaoke,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M., Sumi.N., Ishii.Y., Nakamura,S., Hazama.M., Nishine.T., Harada,A., Yamamoto,R., Matsumoto.H., Sakaguchi.S., Ikegami.T., Kasaiwaqi.K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa.M., Ohara.E., Watabiki.M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuyura,S., Kawal.J., Okazaki,Y., Muramatsu,M., Inoue,K., Kita.A. and Hayashizaki,Y.
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL MEDLINE	Genome Res. 10 (11), 1757-1771 (2000) 20530913
PUBMED	11076661
REFERENCE	4 (sites)
AUTHORS	The RIKEN Genome Exploration Research Group Phase II Team and the PANTOM Consortium.
TITLE	Functional annotation of a full-length mouse cDNA collection
JOURNAL MEDLINE	Nature 409, 685-690 (2001) 5 (bases 1 to 1114)
PUBMED	11076661
REFERENCE	Aachidi,U., Aizawa,K., Akita,H., S., Akimura,T., Aono,H., Arai,A., Atakawa,T., Baldarelli,R., Bono,H., Brownstein,M., Bull,C., Carninci,P., Fukuda,S., Fukushima,Y., Furumori,M., Hanagaki,T., Hara,A., Hayatsu,N., Hill,D., Hiramoto,K., Hirooka,T., Horii,F., Hume.D., Imotani,K., Ishii,Y., Itoh,M., Izawa.M., Kasukawa,T., Kato,H., Kawasaki,J., Kojima,Y., Komoh.H., Kouda.M., Koya.S., Kurihara.C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K., Numasaki.R., Oho.M., Okazaki,Y., Okido,T., Owa.C., Quackenbush,J., Saito,H., Saito,R., Sakai,C., Sakai,K., Sanoh,H., Sasaki,D., Schiraldi,L., Shibata,K., Shihata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi.F., Tanaka,T., Tejima,T., Toya,T., Tamamura,T., YamanaKa,I., Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
TITLE	Direct Submission
JOURNAL	Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
COMMENT	Please visit our web site (http://genome.gsc.riken.go.jp/) for further details. cDNA library was prepared and sequenced in Mouse Genome

Enyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGGAGACGATCCAGACCTCTTTTCTTTTTTCTTCTCAAGAAGTGGA 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 7.5 and subtraction to Rot = 37.5. Second strand cDNA was prepared with the primer adapter of sequence 15' GAGAGGAGATTCGACGTAAATTAAATAATCCCCCCCCC 3'. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI. Host: SOLR.

FEATURES

source

Location/Qualifiers

1..1114

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="MGD:MGI:1893659"

/db_xref="taxon:10090"

/clone="261005N2"

/clone_lib="RIKEN full-length enriched mouse cDNA library"

/dev_stage="10 days embryo"

1..1114

/gene="Rnf7"

1..1114

/gene="Rnf7"

/note="data source:MCD, source key:MGI:1337096, evidence:ISS

ring finger protein 7"

BASE COUNT

289 a 220 c 287 g 318 t

ORIGIN

Query Match

91.3%: Score 1041; DB 11; Length 1114;

Best Local Similarity 99.3%: Pred. No. 3.7e-227;

Matches 1109; Conservative 0; Mismatches 0; Indels 8; Gaps 6;

OY

4

CTGGCGCCGCCGCATATGGCCGACGTGGAGACGGCGAGAACCCCTGCTCTTCTTTCACA

63

Db

2

CTGGCCGCCGCCGCATATGGCCGACGTGGAGAGA-GCGGAGAACCCCTGCTCTTCTTTCACA

60

OY

64

CTCCGGGAGCGCAGGCTCCAAAGTCGGGAGCGCACAAATGTTCTCTCTCAAGAACTGGAA

123

Db

61

CTCCGGGAGCGC-GGCTCCAAAGTCGGGAGCGCACAAAGTGTCTCTCTCAAGAACTGGAA

119

OY

124

GCGCGTAGCCAAATGTGGACTGGGAGCTTGAGTGCATACCTGTGCCATCTGAGGGTCCA

183

Db

120

GCGGTAACCATGTGGACTGGGAGCTTGAGTGCATACCTGTGCCATCTGAGGGTCCA

179

OY

184

GGTGATGATGCTGCTCCCTTCATGTCAAGCTGAAAACAAGCAAGAGCATGTGTGTGGT

243

Db

180

GSTGATGATGCTGCTCCCTTCATGTCAAGCTGAAAACAAGCAAGAGCATGTGTGTGGT

239

OY

244

CTGGGAGAGTGTAACCATTCCTTCACAACTGCTGCATGTCTCTGTGGGTGAACAGAA

303

Db

240

CTGGGAGAGTGTAACCATTCCTTCACAACTGCTGCATGTCTCTGTGGGTGAACAGAA

299

OY

304

CAATGCGTGCCTCTGTGCGACGACGACTGGGTATCTCAAAGAAATCGGCAATGAGAGT

363

Db

300

CAATGCTGCCCCCTGTGTGCGACGACGACTGGGTATCTCAAAGAAATCGGCAATGAGAGT

359

OY

364

GGCCGAGCGCTCTCTGTGTGTGGTTGCTCAACCTTGAGACAAGACTAAGCACTGCAGGGAT

423

Db

360

GGCCGAGCGCTCTCTGTGTGTGGTTGCTCAACCTTGAGACAAGACTAAGCACTGCAGGGAT

419

OY

424

TCAATCTTGAGAGAGAGATGCTGTGCGCTTGTGAAGACTTCACCAAGAGCTTCTTAT

483

Db

420

TCAATCTTGAGAGAGAGATGCTGTGCGCTTGTGAAGACTTCACCAAGAGCTTCTTAT

479

OY

484

TAAATTTGCTGTTTGTGTTTGGGAAATCTCACAATTAAGTAAATTTGTTAAAAATGGC

543

Db

480

TAAATTTGCTGTTTGTGTTTGGGAAATCTCACAATTAAGTAAATTTGTTAAAAATGGC

539

QY	544	CTTTCCACACCTCGTGTGTGTGTGTATACGATGTATGAAAGAGAGAACCCAGAA	603
Db	540	CTTTCCACACCTCGTGTGTGTGTGTATACGATGTATGAAAGAGAGAACCCAGAA	599
QY	604	AATGATCTTTTATCTGTATACCAACGACT--GGAACATTGTGTTACAGAACACATT	660
Db	600	AATGATCTTTTATCTGTATACCAACGACTGTGGGAACATTGTGTACAGAACACATT	659
QY	661	GTTTGTGTTTATGCTGTCGAGGCTTAAAAATATGATTAACGATGTACAGTAACAAATAA	720
Db	660	GTTTGTGTTTATGCTGTCGAGGCTTAAAAATATGATTAACGATGTACAGTAACAAATAA	719
QY	721	ATGCAATGAAAAAGCCGACTCCTCTCTAATCCTTTTGTGTGTGGAGAGAGCCAGCAGCC	780
Db	720	ATGCAATGAAAAAGCCGACTCCTCTCTAATCCTTTTGTGTGTGGAGAGAGCCAGCAGCC	779
QY	781	CACCGTGTGCTTCATTTGTGTGTGAATGAGAGATTTTAACTGACACAGTGAAGAGAGC	840
Db	780	CACCGTGTGCTTCATTTGTGTGTGAATGAGAGATTTTAACTGACACAGTGAAGAGAGC	839
QY	841	TAACTGTGCGGTAAACCTGTAATATGCGCTAACTGTGCGGTAAAGCGCTTGTCTCTGAC	900
Db	840	TAACTGTGCGGTAAACCTGTAATATGCGCTAACTGTGCGGTAAAGCGCTTGTCTCTGAC	899
QY	901	TTTTCACATCTTTGACTGTGGCCAGAGAGCCGTGATTTGTTCAACCACTTGTCTTAAAGAC	960
Db	900	TTTTCACATCTTTGACTGTGGCCAGAGAGCCGTGATTTGTTCAACCACTTGTCTTAAAGAC	959
QY	961	TGTTTCTGTTTTCGCCAAGGTTTATTTGATGTTTATGACAAATAATTTAGTAGAAA	1020
Db	960	TGTTTCTGTTTTCGCCAAGGTTTATTTGATGTTTATGACAAATAATTTAGTAGAAA	1017
QY	1021	ATGCGTTACATATATACACTGTAAGTCAATTTATGCAATGTTTAAATAATATGTGCTT	1080
Db	1018	AT-GCTTACTAGTATTAACACGTAAGTCAATTTATGCAATGTTTAAATAATATGTGCTT	1076
QY	1081	TGAGTTATTAAGTTGATATATATCTCTTAAATCAT	1117
Db	1077	TGAGTTATTAAGTTGATATATATCTCTTAAATCAT	1113
RESULT	3		
LOCUS	AK003248	1084 bp	mRNA linear HTC 19-JAN-2002
DEFINITION	AK003248	Mus musculus 18 days embryo whole body cDNA, RIKEN full-length enriched library, clone:1110001M08:ring finger protein 7, full insert sequence.	
ACCESSION	AK003248		
VERSION	AK003248.1	GI:12833795	
KEYWORDS	HTC; CAP trapper.		
SOURCE	Mus musculus (strain:C57BL/6J) 18 days embryo cDNA to mRNA, clone:111001M08.		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.		
AUTHORS	1 (sites)		
TITLE	Carninci, P., and Hayashizaki, Y.		
JOURNAL	High efficiency full-length cDNA cloning		
MEDLINE	Meth. Enzymol. 303, 19-44 (1999)		
PUBMED	99279253		
REFERENCE	10349636		
AUTHORS	2 (sites)		
TITLE	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.		
JOURNAL	Normalization and subtractions of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes		
MEDLINE	Genome Res. 10 (10), 1617-1630 (2000)		
PUBMED	20499374		
REFERENCE	11042159		
AUTHORS	3 (sites)		
TITLE	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Kitsuana, T., Tashiro, H., Itoh, M.,		

TITLE	JOURNAL	MEDLINE	PUBMED	REFERENCE	AUTHORS
THE RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium Functional annotation of a full-length mouse cDNA collection Nature 409, 685-690 (2001) 5 (bases 1 to 1084)					Adachi,T., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A., Arakawa,T., Baldarelli,R., Bono,H., Brownstein,M., Bulic,C., Carninci,P., Fukuda,S., Fukushima,Y., Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hill,D., Hiramoto,K., Hirooka,T., Horii,F., Hume,D., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T., Kato,H., Kawaji,J., Kojima,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K., Numasaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Quackenbush,J., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Schirml,L., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yamana,K.I., Yasunishii,A., Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
TITLE	JOURNAL	MEDLINE	PUBMED	REFERENCE	AUTHORS
cdna library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cdna was primed with a primer [5' GAGGAGACAGCGCCGCCAATCTGGTGTGTTTGTGTA 3']. cdna was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cdna went through one round of normalization to kot = 5.0. Second strand cdn was prepared with the primer adapter of sequence[5' GAAGAAGAGAGATCCAGACGTCATCAATTATTAATAAACCCCCCCC 3']. cdna was cloned with XhoI and SstI. Cloning sites, 5' end: SstII; 3' end: XhoI. Host: SOLR.					
FEATURES	SOURCE				
gene	CDS				
location/Qualifiers					
1..1084	/organism="Mus musculus"				
	/strain="C57BL/6J"				
	/db_xref="MGI:1892372"				
	/db_xref="taxon:10090"				
	/clone="1110001M08"				
	/clone_lib="RIKEN full-length enriched mouse cdna library"				
	/dev_stage="18 days embryo"				
	15..356				
	/gene="Rnf7"				
	15..356				
	/gene="Rnf7"				
	/note="data source:MGI, source key:MGI:1337096, evidence:ISS				
	putative				
	ring finger protein 7"				
	/codon_start=1				
	/protein_id="BAB2666.1"				
	/db_xref="GI:12833796"				

/translation="MADVEDEEPCVLSHSGSAGSKSGDKMFSLKKNVAVAMNSMD
 VEDPCICRVOVMDACLRQCAENKQEDCVVWVWGEONHNFHNCMSIMWKQKNRCPIC
 ODDMYVORIGK"
 polyA_signal 1062..1067
 /note="putative"
 polyA_site 1084
 /note="putative"
 BASE COUNT 277 a 218 c 285 g 304 t
 ORIGIN

Query Match 90.6%; Score 1032.4; DB 11; Length 1084;
 Best local similarity 99.4%; Pred. No. 3.4e-225;
 Matches 1079; Conservative 0; Mismatches 1; Indels 6; Gaps 4;

QY 4 CTCGCGCGCGCGCATGGCCGACGTGGAGACGGCGAGAACCTCGCTCTTCTTCGCA 63
 |||||||
 DB 2 CTCGCGCGCGCGCATGGCCGACGTGGAGACGGCGAGAACCTCGCTCTTCTTCGCA 61
 |||||||
 QY 64 CTCGCGAGGCGCAGGCTCCAAAGTCGGAGAGGCAAGATGTCTCTCTCAAGAATGGAA 123
 |||||||
 DB 62 CTCGCGAGGCGCAGGCTCCAAAGTCGGAGAGGCAAGATGTCTCTCTCAAGAATGGAA 121
 |||||||
 QY 124 CGCGGTAAGCCATGTGAGCTGGAGACGTTGAGTCGATACCTGTGCCATCTGCAGGGTCA 183
 |||||||
 DB 122 CGCGGTAAGCCATGTGAGCTGGAGACGTTGAGTCGATACCTGTGCCATCTGCAGGGTCA 181
 |||||||
 QY 184 GGTGATGATGCTGCTGCTGTGATGTCAAGTGAAGAACAGAGAGAGCTGTGTGTGT 243
 |||||||
 DB 182 GGTGATGATGCTGCTGCTGTGATGTCAAGTGAAGAACAGAGAGAGCTGTGTGTGT 241
 |||||||
 QY 244 CTGGGAGAGTGTACCATCTCTCCCAACTGCTGCATGCTCCGTGGGTGAAGACGAA 303
 |||||||
 DB 242 CTGGGAGAGTGTACCATCTCTCCCAACTGCTGCATGCTCCGTGGGTGAAGACGAA 301
 |||||||
 QY 304 CAATCGCTGCTCTGTGTGCCAGCAGAGACTGGTATGCCAAGAATCGCAATGAGAGGT 363
 |||||||
 DB 302 CAATCGCTGCTCTGTGTGCCAGCAGAGACTGGTATGCCAAGAATCGCAATGAGAGGT 361
 |||||||
 QY 364 GGGCCAGGCGCTCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 423
 |||||||
 DB 362 GGGCCAGGCGCTCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 421
 |||||||
 QY 424 TCATCTCTGAGAGAGAGAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 483
 |||||||
 DB 422 TCATCTCTGAGAGAGAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 481
 |||||||
 QY 484 TAAATGTCTGT 543
 |||||||
 DB 482 TAAATGTCTGT 541
 |||||||
 QY 544 CTTTCTCTACTCTGCTGT 603
 |||||||
 DB 542 CTTTCTCTACTCTGCTGT 601
 |||||||
 QY 604 AATGATCTTGT 660
 |||||||
 DB 602 AATGATCTTGT 661
 |||||||
 QY 661 GTTTGT 720
 |||||||
 DB 662 GTTTGT 721
 |||||||
 QY 721 ATGCATTTGAAG 780
 |||||||
 DB 722 ATGCATTTGAAG 781
 |||||||
 QY 781 CACCTGCTGTCTTCAATTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 840
 |||||||
 DB 782 CACCTGCTGTCTTCAATTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 841
 |||||||
 QY 841 TAACTGTGGGTAACTGTATATATGCGGTAACTGTGGGTAAAGCGCTTGTCTCTGTAC 900
 |||||||

DB 842 TAACTGTGGGTAACTGTATATATGCGGTAACTGTGGGTAAAGCGCTTGTCTCTGTAC 901
 |||||||
 QY 901 TTCTCCATCTTTTACTTTGGCCAGAGACCTGTGATTTGTCAACCACTTACTTCAAGAAC 960
 |||||||
 DB 902 TTCTCCATCTTTTACTTTGGCCAGAGACCTGTGATTTGTCAACCACTTACTTCAAGAAC 961
 |||||||
 QY 961 TGTTCCTGTTTTTTGGCGAAGGTTGATTTGATTTTACTCAAAAATATTTAGTAGAAA 1020
 |||||||
 DB 962 TGTTCCTGTTTTTTGGCGAAGGTTGATTTGATTTTACTCAAAAATATTTAGTAGAAA 1019
 |||||||
 QY 1021 ATGCTTACTAGTATTAACACTGAAGTTCATATGCAATGTTTAAATATTTGCTT 1080
 |||||||
 DB 1020 AT-GCTTACTAGTATTAACACTGAAGTTCATATGCAATGTTTAAATATTTGCTT 1078
 |||||||
 QY 1081 TGAGTT 1086
 |||||||
 DB 1079 TGAGCT 1084

RESULT 4

AK002628

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

TITILE

JOURNAL

MEDLINE

PUBMED

AUTHORS

REFERENCE

AUTHORS

TITILE

JOURNAL

MEDLINE

PUBMED

AUTHORS

REFERENCE

AUTHORS

TITILE

JOURNAL

MEDLINE

PUBMED

AUTHORS

REFERENCE

AK002628 1078 bp mRNA linear HTC 19-JAN-2002
 Mus musculus adult male kidney cDNA, RIKEN full-length enriched
 library, clone:0610013021:ring finger protein 7, full insert
 sequence.

AK002628.1 GI:12832752
 HTC; CAP trapper.
 Mus musculus (strain:C57BL/6J) adult male kidney cDNA to mRNA,
 clone:lib:RIKEN full-length enriched mouse cDNA library
 clone:0610013021.
 Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (sites)

Carninci, P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning
 Meth. Enzymol. 303, 19-44 (1999)

99279253
 10349636

2 (sites)
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 Genome Res. 10 (10), 1617-1630 (2000)

JOURNAL
 MEDLINE
 PUBMED
 20493374
 11042159

3 (sites)
 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
 Konno, H., Akiyama, J., Nishi, K., Kitsuina, T., Tashiro, H., Itoh, M.,
 Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Washiwa, K.,
 Fujiwake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Kashiwagi, M.,
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system-384-format
 sequencing pipeline with 384 multicapillary sequencer
 Genome Res. 10 (11), 1757-1771 (2000)

JOURNAL
 MEDLINE
 PUBMED
 20530913
 11076861

4 (sites)
 The RIKEN genome Exploration Research Group Phase II Team and the
 FANTOM Consortium.
 Functional annotation of a full-length mouse cDNA collection
 Nature 409, 685-690 (2001)

5 (bases 1 to 1078)
 Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A.,
 Arakawa, T., Baldairelli, R., Bono, H., Brownstein, M., Bull, C.,
 Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T.,
 Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hirooka, T., Hoti, F.,
 Hume, D., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T.,
 Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S.,

Technologies, catalog #12017-018. Investigators providing samples: Lothar Hennigshausen/Chu-Xia Deng, NIH Reference for transgenic model: Xu et al., Nature Genetics 22, 37-43 (1999). Note: this is a NCI-CCAP Library."

BASE COUNT	250 a	199 c	260 g	269 t
ORIGIN				

Query Match	59.1%;	Score 673.4;	DB 10;	Length 978;
Best Local Similarity	90.5%;	Pred. No. 2.8e-143;		
Matches 837;	Conservative	0;	Mismatches 61;	Indels 27;
				Gaps 10;

Db	421	AGGATCGCTGGGCCCTTTGAAGCTACCAAGAAGCCTTGCTTAATTAAATTGTGTGTTACT	480
OY	501	TTTTGGAAATTCCTCACATTTAGATAATTTGTTAAAAATGGCCCTTCTTACCCTGTGTG	560
Db	481	TTTGGGAAAATTCCTCACAAATTAAGATAATTTGTTAAAAATGGCCCTTCTTACCCTGTGTG	540
OY	561	TGTGTGTGTGATACAAATGCATAGAAAGCGAAGAACACACAGAAATGATCTTGTGTTATC	620
Db	541	TGTGTGTGTGATACAAATGCATAGAAAGCGAAGAACACACAGAAATGATCTTGTGTTATC	600
OY	621	TGTACCCACGACT--GGACAATTGTGTTCACAGAAATCATTTGTTGTTATGCTTG	677
Db	601	TGTACCCACGACTGTGGGAACTTGTGTTCACAGAAACAA-TGTACAGTGTATAGCTTG	659
OY	678	AGGGTAAAAAATATATATAAGATTAAGATGTTACACTTAACAATATAATGATGAAAGCCGA	737
Db	660	AGGGTAAAAAATATATATAAGATTAAGACACTGTTACAGTATAC-AATATAATGATGAAAGCCGA	718
OY	738	CTCCTCCTAATCC	750
Db	719	CTCCTCCTAATCC	731
RESULT	9		
LOCUS	B1156184	978 bp	mRNA linear EST 05-JUL-2001
DEFINITION	602903365F1 NIH_CGAP_Mam3 Mus musculus cDNA clone IMAGE:5032984 5'		
ACCESSION	B1156184		
VERSION	B1156184.1	GI:14616185	
KEYWORDS	EST.		
SOURCE	house mouse.		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
TITLE	1 (bases 1 to 978)		
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/.		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC) unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgapbs-remail.nih.gov		
	Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.		
	cDNA Library Preparation: Life Technologies, Inc.		
	cDNA library Arrayed by: The I.M.A.G.E. Consortium (LMN)		
	DNA Sequencing by: Incyte Genomics, Inc.		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LMN at:		
	http://image.llnl.gov		
	Plate: L1AM11092 row: a column: 17		
	High quality sequence stop: 855.		
FEATURES			
Source	Location/Qualifiers		
	1..978		
	/organism="Mus musculus"		
	/db_xref="taxon:10090"		
	/clone="IMAGE:5032984"		
	/clone_lib="NIH_CGAP_Mam3"		
	/tissue_type="tumor, gross tissue"		
	/lab_host="DH10B"		
	/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: NotI;		
	Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.		
	Average insert 2 kb. Library constructed by Life		

OY	75	CAGGCTCAAGTCGGGAGGCGCAGAGATGTCTCTCTCAAGAAAGTGGAAACGGCGTAGCCA	134
Db	1	CAGGCTCCAGATCGGSGAGGCCCAAGATGTTCTCTCTCAAGAAAGTGGAAACGGCGTAGCCA	60
OY	135	TGTGAGCTGGAGCTGTAGTGTGATGTACCTGTGTCATCTGCACGGGTCCAGTGTATGATG	194
Db	61	TGTGAGCTGGAGCTGTAGTGTGATGTACCTGTGTCATCTGCACGGGTCCAGTGTATGATG	120
OY	195	CCTGCTTCGATGTCAAGCTGAAAACAGCAGAGAGACTGTGTGGCTGCGGGAGAGT	254
Db	121	CCTGCTTCGATGTCAAGCTGAAAACAGCAGAGAGACTGTGTGGCTGCGGGAGAGT	180
OY	255	GTAACCATTCCTCCACAACACTGCTCATGTCCCTGTGGGTGGAACAGAAATTCGTGCC	314
Db	181	GTAACCATTCCTCCACAACACTGCTCATGTCCCTGTGGGTGGAACAGAAATTCGTGCC	240
OY	315	CTCTGTGCGCAGAGAGACTGGGTATGTCCAAAGAAATGGGCAATGAGAGGGTGGCCACGGCC	374
Db	241	CTCTGTGCGCAGAGAGACTGGGTATGTCCAAAGAAATGAGAGGGTGGCCACAGGCC	300
OY	375	TCTCTGTGTGTGTGCTGACCCCTGACAAAGACTTAAACACTGCAAGGGGATTCATCCTTGAG	434
Db	301	TCTCTGTGTGTGTGCTGACCCCTGACAAAGACTTAAACACTGCAAGGGGATTCATCCTTGAG	360
OY	435	AGAGAGAGATTCCTGTGGCCCTTTAGACTCACCAGGCTTGTTATTAATTTGTCTG	494
Db	361	AGAGAGAGATTCCTGTGGCCCTTTAGACTCACCAGGCTTGTTATTAATTTGTCTG	420
OY	495	TTTACTTTTGGGAAATTCCTCTACAAATTAAGATAATTTGTTAAAAATGCGCTTCTCTACT	554
Db	421	TTTACTTTTGGGAAATTCCTCTACAAATTAAGATAATTTGTTAAAAATGCGCTTCTCTACT	480
OY	555	CTGTGTGTGTGTGTGATATACGAATGCATATGAAAGAGCGAGAACACCGAAGAAATGATCTTGG	614
Db	481	CTGTGTGTGTGTGTGATATACGAATGCATATGAAAGAGCGAGAACACCGAAGAAATGATCTTGG	540
OY	615	TTTATCTGTACCCAGCAGACT---GGAACATTTGTTTACAGAGAACATTTGTTGTTT	671
Db	541	TTTATCTGTACCCAGCAGACTGTGGAAACATTTGTTTACAGAGAACATTTGTTGTTT	600
OY	672	TGCTTGAGGGTTAAAAATAGATTAACGATG--TTACAGTAAACAATTAATATGATGAA	730
Db	601	TGCTTGAGGGTTAAAAATAGATTAACGATG--TTACAGTAAACAATTAATTAATATGATGAA	660
OY	731	AAGCGAGCTCCCTCCAAATCCCTTTTGTGTGGGAGAGAGCAGCGAG---GCCACC	785
Db	661	AAGCGAGCTCCCTCCAAATCCCTTTTGTGTGGGAGAGAGCGAAGCGAAGGCCACCCCT	720
OY	786	TGCTGTCTTTCATTTCTCT-----GTGAATGAGGATTTTAACTGTCACTCAGTGAAG--AGGC	839
Db	721	TGCTGTCTTTCATTTCTCTGTGTGAATGAGCGGATTTTAACTGTCACTCAGTGAAGAGCC	780
OY	840	GTAAGTGTGCGG--TAAACTGTAAATGTGCGGTAA-----CTGTGCGGTAAACGGCTTTGTC	893
Db	781	GTAAGTGTGCGGGTAAACTGTAAATGTGCGGTAAACCTGTGCGGTAAACGGCTTTGTC	840
OY	894	T--CCGAGCTTCCACATCTT---GACTGGGCGAGAGACC--TGATTTTTCGAACCATCT	947
Db	841	TCCCTGACTTTCACATCTTTCGAACCTTGTGGCGAGAGAACCTTGGGATTTGTCAACCATCT	900
OY	948	AGTTCTAAAGACTGTTTCTGTTT	972

```

Db      901 AGTTCTAAAGCAACGGTCTGTT 925
|||||
RESULT 10
LOCUS   BG917948          916 bp    mRNA          linear    EST 05-JUN-2001
DEFINITION 602820877F1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:4949851 5',
            mRNA sequence.
ACCESSION BG917948
VERSION   BG917948.1 GI:14298424
KEYWORDS  EST.
SOURCE    house mouse.
ORGANISM  Mus musculus
            Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 916)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgabds-remail.nih.gov
            Tissue Procurement: Jeffrey Green M.D.
            CDNA Library Preparation: Life Technologies, Inc.
            DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10903 row: 1 column: 20
            High quality sequence stop: 760.
FEATURES
    source
        1..916
            /organism="Mus musculus"
            /strain="FVB/N"
            /db_xref="taxon:10090"
            /clone="IMAGE:4949851"
            /clone_lib="NCI_CGAP_Mam6"
            /sex="female, virgin"
            /tissue_type="infiltrating ductal carcinoma"
            /dev_stage="5 months"
            /lab_host="DH10B"
            /note="Organ: mammary; Vector: pCMV-SPORT6; Site: 1: SalI;
            Site: 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
            Library constructed by Life Technologies. Investigator
            providing samples: Jeffrey Green, M.D., NIH"
BASE COUNT 232 a 202 c 264 g 218 t
ORIGIN
Query Match      58.4%; Score 665.4; DB 10; Length 916;
Best Local Similarity 93.8%; Pred. No. 1.9e-141;
Matches 761; Conservative 0; Mismatches 36; Indels 14; Gaps 6;
QY      1 GTTCTGGCGCCGCGCATGGCCGACGTGAGAGCGGCGAGAACCTGCTCTTCTTC 60
        |||||||
Db      40 GTTCTGGCGCGCCGCGCATGGCCGACGTGAGAGCGGCGAGAACCTGCTCTTCTTC 99
QY      61 GCACCTCGGAGGCGGACGCTCCAAGTGGGAGCGGACACATGTTCTCTCAAGAGTG 120
        |||||||
Db      100 GCACCTCGGAGGCGGACGCTCCAAGTGGGAGCGGACACATGTTCTCTCAAGAGTG 159
QY      121 GAACGGCGGTAGCATGTGAGAGCTGGGAGCTTGAGTGCGATACCTGTCCTGCGAGGGT 180
        |||||||
Db      160 GAACGGCGGTAGCATGTGAGAGCTGGGAGCTTGAGTGCGATACCTGTCCTGCGAGGGT 219
QY      181 CCAAGTGATGATGCTGCTGCTGATGTCAGGTGAAGAACACAGAGAGACTGTGTTGT 240
        |||||||
Db      220 CCAAGTGATGATGCTGCTGCTGATGTCAGGTGAAGAACACAGAGAGACTGTGTTGT 279
QY      241 GGTCTGGGAGAGTGTACCATTCCTTCCCAAACTGCTGATGTCCTGCTGGGTGAACA 300
        |||||||
Db      280 GGTCTGGGAGAGTGTACCATTCCTTCCCAAACTGCTGATGTCCTGCTGGGTGAACA 339

```

```

QY      301 GAACAATCGCTGCGCTCTGTGAGCAGAGACCTGGGTAGTCCAAAGATGGCAATGAGA 360
        |||||||
Db      340 GAACAATCGCTGCGCTCTGTGAGCAGAGACCTGGGTAGTCCAAAGATGGCAATGAGA 399
QY      361 GGTGGCCCAAGGCGCTCTGTGAGTGGTGGTGCAGACCTGGACAAAGACTGACGAGG 420
        |||||||
Db      400 GGTGGCCCAAGGCGCTCTGTGAGTGGTGGTGCAGACCTGGACAAAGACTGACGAGG 459
QY      421 GATTCAATCTTGAGAGAGAGAGATGCTGCGCCTTTGAGACTCACAAAGCCTTGCTT 480
        |||||||
Db      460 GATTCAATCTTGAGAGAGAGAGATGCTGCGCCTTTGAGACTCACAAAGCCTTGCTT 519
QY      481 TATTAAATTTGCTGTTAGTGTGGGAAATTCCTACATTAAGATTAATTTGTTAAAT 540
        |||||||
Db      520 TATTAAATTTGCTGTTAGTGTGGGAAATTCCTACATTAAGATTAATTTGTTAAAT 579
QY      541 GGCCTTCTTACCTGCTGAGTGTGTGTGATGATGATGATGATGATGATGATGATGATG 600
        |||||||
Db      580 GGCCTTCTTACCTGCTGAGTGTGTGTGATGATGATGATGATGATGATGATGATGATG 638
QY      601 GAAATGATCTTTGTTATCTGTACCCAGAGCT--GGAACATTTGTTACAGAGAAC 657
        |||||||
Db      639 GAAATGATCTTTGTTATCTGTACCCAGAGCTGAGGAAATTTGTTACAGAGAAC 698
QY      658 ATGTTTGTGTTATCTGTAGAGGTT-AAAAATAGATTAAGATTTAGTACAGTACAA 716
        |||||||
Db      699 CATG--TGCGGTATCTGTGAGGTTAAATAATAGATTAAGATTTAGTACAGTACAA 756
QY      717 TAAAT--TGCAATGAAAGCCGACTGCTCTTAATCTTTTGTGT--TGAGAGAGAG 769
        |||||||
Db      757 ATAAATATGCGCTTGAAGAGCCGAAATCTCTTAATGCTTGTAGTGTAGGAGAGAG 816
QY      770 GCAAGCAGGCGCCCTGCTGTTCAATTTG 800
        |||||||
Db      817 GCAAGCAGGAGCAACCTGTTCTTCATTTG 847

```

RESULT 11
B1147544
LOCUS B1147544 754 bp mRNA linear EST 05-JUL-2001
DEFINITION 602913971F1 NCI_CGAP_L19 Mus musculus cDNA clone IMAGE:5055160 5',
mRNA sequence.
ACCESSION B1147544
VERSION B1147544.1 GI:14607545
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 754)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabds-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11149 row: m column: 17
High quality sequence stop: 665.
FEATURES
source
1..754
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:5055160"
/clone_lib="NCI_CGAP_L19"
/lab_host="DH10B (T1 phage-resistant)"

B1556918
 LOCUS 737 bp mRNA linear EST 05-SEP-2001
 DEFINITION 603239970F1 NIH_CGAP_Mam3 Mus musculus cDNA clone IMAGE:5292832 5', mRNA sequence.
 ACCESSION B1556918
 VERSION B1556918.1 GI:15444232
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 737)
 NIH-MGC <http://mgc.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC).
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Plate: L1AM11740 row: 1 column: 17
 High quality sequence stop: 685.
 Location/Qualifiers
 1..737
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone="IMAGE:5292832"
 /clone_1lb="NIH_CGAP_Mam3"
 /tissue_type="tumor, gross tissue"
 /lab_host="DH10B"
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site:1: NotI; Site:2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert 2 kb. Library constructed by Life Technologies, catalog #12017-018. Investigators providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH Reference for transgenic models: Xu et al., Nature Genetics 22, 37-43 (1999). Note: this is a NCI_CGAP Library."
 BASE COUNT 200 a 152 c 212 g 173 t
 ORIGIN

Query Match 56.7%; Score 646.2; DB 10; Length 737;
 Best Local Similarity 97.1%; Pred. No. 4.4e-137;
 Matches 702; Conservative 0; Mismatches 13; Indels 8; Gaps 4;

QY 20 GCCGACGTGAGAGAGCGGAGAACCTGCTCTTCTTCGACCTCCGGAGCGCAGGC 79
 |||||||
 Db 1 GCCGACGTGAGAGAGCGGAGAACCTGCTCTTCTTCGACCTCCGGAGCGCAGGC 60
 |||||||

QY 80 TCCAAGTGGGAGGCGACAAAGATGTTCTCTCAAGAAGTGGAGCGCGTACCATGTGG 139
 |||||||
 Db 61 TCCAAGTGGGAGGCGACAAAGATGTTCTCTCAAGAAGTGGAGCGCGTACCATGTGG 120
 |||||||

QY 140 AGCTGGAGCGTTGAGTGGGATACCTGTCATCGACAGGTCCAGAGTATGATGCTCGC 199
 |||||||
 Db 121 AGCTGGAGCGTTGAGTGGGATACCTGTCATCGACAGGTCCAGAGTATGATGCTCGC 180
 |||||||

QY 200 CTTGCGATGTCAGCTGAAGAACAGACAGAGACTGTGTGTGCTGGGAGAGCTGTAAC 259
 |||||||
 Db 181 CTTGCGATGTCAGCTGAAGAACAGACAGAGACTGTGTGTGCTGGGAGAGCTGTAAC 240
 |||||||

QY 260 CATTCCTTCACAACTGCTGATGCCCTGTGGGGAACAAGCAATCGCTGCCCTGTG 319
 |||||||
 Db 241 CATTCCTTCACAACTGCTGATGCCCTGTGGGGAACAAGCAATCGCTGCCCTGTG 300
 |||||||

QY 320 TGCCAGCAGAGCTGGGTAGTCCAAAGATCGCAATGAGAGTGGCCAGGCGCTCCTG 379
 |||||||
 Db 301 TGCCAGCAGAGCTGGGTAGTCCAAAGATCGCAATGAGAGTGGCCAGGCGCTCCTG 360
 |||||||

QY 380 GTGTGGTGTGCTGACCTGGACAAAGACTAAAGACTGACAGGGATTTCATCTTGAGAGACA 439
 |||||||
 Db 361 GTGTGGTGTGCTGACCTGGACAAAGACTAAAGACTGACAGGGATTTCATCTTGAGAGACA 420
 |||||||

QY 440 GAGGATGCTGTGCGCCTTTGAGACTCACCAGGAGCTTGTATTTATTTGCTGTTTG 499
 |||||||
 Db 421 GAGGATGCTGTGCGCCTTTGAGACTCACCAGGAGCTTGTATTTATTTGCTGTTTG 480
 |||||||

QY 500 TTTTGGGAAATTCCTACAAATTAAGTAAATTTGTT-AAAAAGGCGCTTCTACCCCTGG 558
 |||||||
 Db 481 TTTTGGGAAATTCCTACAAATTAAGTAAATTTGTT-AAAAAGGCGCTTCTACCCCTGG 540
 |||||||

QY 559 TGTGTGTGTGTGATACGAATGCATAGAGAGCGAGAACCCAGAAATGATCTTTGTTTA 618
 |||||||
 Db 541 TGTGTGTGTGTGATACGAATGCATAGAGAGCGAGAACCCAGAAATGATCTTTGTTTA 600
 |||||||

QY 619 TCTGTACCCACGACT--GGAACATTGTGTTACAGAGAACATTGTTGTGTATGCT 675
 |||||||
 Db 601 TCTGTACCCACGAGTGTGGACATGATGTGTTACAGAGAACATG--AGGTAATGCT 657
 |||||||

QY 676 TGAGGTTAAATAATGATTAACGATTTACAGTAACAATAAA-TCGATTGAAAACG 734
 |||||||
 Db 658 TGAGGTTAAATAATGATTAACGATTTACAGTAACAATAAACTGATTGAAAACG 717
 |||||||

QY 735 CGA 737
 |||
 Db 718 CGA 720
 |||

RESULT 15
 BF302267 822 bp mRNA linear EST 21-NOV-2000
 LOCUS 602031233F1 NCI_CGAP_SG2 Mus musculus cDNA clone IMAGE:416288 5', mRNA sequence.
 DEFINITION BF302267
 ACCESSION BF302267
 VERSION BF302267
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 822)
 NIH-MGC <http://mgc.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC).
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Jeffrey E. Green, M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Plate: L1AM9455 row: a column: 09
 High quality sequence stop: 738.
 Location/Qualifiers
 1..822
 /organism="Mus musculus"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:416288"
 /clone_1lb="NCI_CGAP_SG2"
 /lab_host="DH10B (Y1 phage-resistant)"
 /note="Organ: salivary gland; Vector: pCMV-SPORT6; Site:1: NotI; Site:2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.3 kb. Constructed by Life Technologies. Note: this is a NCI_CGAP Library."
 BASE COUNT 227 a 178 c 229 g 188 t
 ORIGIN

Query Match 54.8%; Score 624.4; DB 10; Length 822;

Best Local Similarity 96.9%; Pred. No. 4.2e-132;
Matches 713; Conservative 0; Mismatches 11; Indels 12; Gaps 7;

```
OY 6 GGGCCGGCCGCGCATGGCCCGACGCGGAGAGAGGAGAGACCCCTGCGCTCTTCTTCTGCGACT 65
    ||| |||||
Db 5 GGGTGCGCGCCATGGCCCGACGCGGAGAGAGGAGAGACCCCTGCGCTCTTCTTCTGCGACT 64
OY 66 CCGGAGCGCAGAGCTCCAGAGTGGGAGGCGACAGATGTTCTCTCTCAAGAGTGAACG 125
    |||||
Db 65 CCGGAGCGCAGAGCTCCAGAGTGGGAGGCGACAGATGTTCTCTCTCAAGAGTGAACG 124
OY 126 CCGTAGCCATGTGAGCTGGGAGCTGTGAGTGGCATACCTGTGCGCATCTGACGGGTCCAGG 185
    |||||
Db 125 CCGTAGCCATGTGAGCTGGGAGCTGTGAGTGGCATACCTGTGCGCATCTGACGGGTCCAGG 184
OY 186 TGATGATGCGCTGCTTGTGATGTCAAGCTGAAGAAACAGCAGAGACTGTGTGTGCT 245
    |||||
Db 185 TGATGATGCGCTGCTTGTGATGTCAAGCTGAAGAAACAGCAGAGACTGTGTGTGCT 244
OY 246 GGGGAGAGTGTAAACCATTCCTTCCACAACCTGCTGCATGTCCCTGTGGGTGAACAGAAC 305
    |||||
Db 245 GGGGAGAGTGTAAACCATTCCTTCCACAACCTGCTGCATGTCCCTGTGGGTGAACAGAAC 304
OY 306 ATCGCTGCCCTGTGTGCCAGCAGAGACTGGGTAGTCCAAAGAAATCGGCAAAATGAGAGGTGG 365
    |||||
Db 305 ATCGCTGCCCTGTGTGCCAGCAGAGACTGGGTAGTCCAAAGAAATCGGCAAAATGAGAGGTGG 364
OY 366 CCCAGGCGCTCCTGTGTGGTGTGCTGACCCCTGGACAAAGACTAAACACTGACGAGGGGATTG 425
    |||||
Db 365 CCCAGGCGCTCCTGTGTGGTGTGCTGACCCCTGGACAAAGACTAAACACTGACGAGGGGATTG 424
OY 426 ATCCTTGAGAGAGAGAGATGCTGTGCGCTTGTGAGACTCACCAAGGCTTCTTATTA 485
    |||||
Db 425 ATCCTTGAGAGAGAGAGATGCTGTGCGCTTGTGAGACTCA--AAAGCTTCTTATTA 482
OY 486 ATTTGCTGTGTTAGTTTGGGAAATTCCTACATTAAGATAATTTGTTAAAAATGGCCT 545
    ||| | | |
Db 483 AT--TGCTGTGTTAGTTTGGGAAATTCCTACATTAAGATAA--TTGTTAAAAATGGCCT 538
OY 546 TTCCACCTCTGCTGTGTGTGTGTGATACGATGATAGAAAGAGAGACACCAAGAAAA 605
    |||||
Db 539 TTCCACCTCTGCTGTGTGTGTGTGATACGATGATAGAAAGAGAGACACCAAGAAAA 598
OY 606 TGATCTTTTATCTGTACCCAGCACT--GGAACATTTGTCTACAGAGAGACATTTGT 662
    |||||
Db 599 TGATCTTTTATCTGTACCCAGCACTGTGGAACATTTGTCTACAGAGAGACA-TGT 657
OY 663 TTGTGTTTATGCTTGAGGGTT--AAAAATAGATTAACGAATGTACAGTAACAATAAAA 721
    |||||
Db 658 TTGTGTTTATGCTTGAGGGTTCAACAAATAGATTAACGAATGTACAGT-ACAAATTA 716
OY 722 TGCATTGAAAAGCCGA 737
    ||| |||||
Db 717 TGCTTTGAAAAGCCGA 732
```

Search completed: July 25, 2002, 18:05:39
Job time: 14218 sec

THIS PAGE BLANK (USPTO)